

TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

COMME EXIGENCE PARTIELLE
À L'OBTENTION DE
MAÎTRISE EN GÉNIE ÉLECTRIQUE

PAR
LEE-PIERRE BELLEY

SÉPARATION DE SOURCES AVEUGLE
UTILISANT LA TRANSFORMÉE EN ONDELETTES
CONTINUES

MONTREAL, LE 14 JANVIER 2008
© droits réservés de Lee-Pierre Belley

CE MÉMOIRE A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE :

M. Christian Gargour, directeur de mémoire
Département de génie électrique à l'École de technologie supérieure

M. Gheorghe Marcel Gabrea, codirecteur de mémoire
Département de génie électrique à l'École de technologie supérieure

M. Jean-Marc Lina, président du jury
Département de génie électrique à l'École de technologie supérieure

M. Maarouf Saad, membre du jury
Département de génie électrique à l'École de technologie supérieure

IL A FAIT L'OBJET D'UNE SOUTENANCE DEVANT JURY ET PUBLIC

LE 20 DÉCEMBRE 2007

À L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

AVANT-PROPOS

La principale raison m'ayant poussé à poursuivre mes études aux niveaux supérieurs fût mon insatisfaction par rapport au niveau de maîtrise des connaissances acquises lors de mon baccalauréat. Ayant toujours cherché à repousser mes propres limites j'ai donc décidé de poursuivre mon éducation.

L'intérêt envers le traitement des signaux provient de mon goût pour la musique. N'étant pas très familier avec le domaine du numérique et celui-ci étant de plus en plus présent dans notre entourage ce choix me semblait très pertinent. Les outils d'enregistrement et de productions sont devenus accessibles à la majorité de la population principalement dû à la diminution des coûts des ordinateurs personnels et l'augmentation de leurs performances. Le besoin de pouvoir développer avec ces plateformes était indéniable.

Ainsi, un sujet de recherche intégrant l'acoustique, l'informatique et les mathématiques me semblait un choix judicieux. Au-delà du sujet qu'est la séparation de source, ce travail de recherche constitue pour moi une formation en traitement numérique des signaux. Elle me permet de comprendre et d'accéder au monde derrière la façade des outils de production numériques.

REMERCIEMENTS

Je tiens à remercier tous ceux qui m'ont soutenu pendant la maîtrise. La motivation et l'objectivité ne sont pas toujours présentes. Tout ceux qui m'entourent m'ont aidé à croire en mes capacités et m'ont donné l'énergie nécessaire pour terminer ce travail. Mon directeur et mon co-directeur ont été deux ressources importantes. Ils m'ont guidé et m'ont permis de comprendre le monde de la recherche et du développement. Je les remercie pour l'aide apportée tout au long de ce travail.

SÉPARATION DE SOURCES AVEUGLE BASÉE SUR L'UTILISATION DE LA TRANSFORMÉE EN ONDELETTES CONTINUES

BELLEY, LEE-PIERRE

RÉSUMÉ

La séparation d'un mélange convolutif de deux sources à partir de deux capteurs est considérée dans ce travail. Plusieurs approches sont présentées dans la littérature qui font usage de méthodes statistiques et de transformées telle que la transformée de Fourier à fenêtre glissante et la transformée en paquets d'ondelettes. Un nouvel algorithme basé sur l'analyse en composantes indépendantes et la transformée en ondelettes continues est présenté. Une méthode connue de séparation de sources dans le domaine de la fréquence a été modifiée en remplaçant l'algorithme d'élimination de la corrélation croisée par un algorithme de séparation dont les performances sont indépendantes de la matrice de mélange. Les coefficients de la transformée en ondelettes continues des mélanges sont employés comme signaux d'entrée dans l'algorithme développé plutôt que les signaux eux mêmes. Les échelles de la transformée sont réparties suivant les schèmes suivants : linéaires, logarithmiques et loi mu. Une distribution uniforme des échelles de fréquence a aussi été considérée. Les mélanges employés pour l'évaluation de l'algorithme sont synthétiques. Les résultats expérimentaux sont présentés et évalués.

Mots clefs : transformée en ondelettes continues, séparation aveugle de sources, analyse en composantes indépendantes, mélanges convolutifs

BLIND SOURCE SEPARATION BASED ON CONTINUOUS WAVELET TRANSFORM

BELLEY, LEE-PIERRE

ABSTRACT

Separation of convolutive mixtures of two speech sources using two sensors is considered in this work. Several approaches have been reported in the literature using statistical methods as well as transforms such as the Short Time Fourier Transform and the Wavelet Packet Transform. We propose a new source separation method based on the independent component analysis and using the Continuous Wavelet Transform. We have modified an existing frequency based separation block by replacing in it the decorrelation algorithm by the Equivariant Adaptive Separation via Independence. We have also used the Continuous Wavelet Transform coefficients as its input rather than the signal itself or its wavelet packet coefficients. The Continuous Wavelet Transform scales repartitions have been chosen according to the following schemes: linear, logarithmic and mu law. A uniform frequency distribution pattern has also been considered. Synthetic mixtures of sources have been used. Experimental results obtained by this method have been presented and evaluated.

Keywords: continuous wavelet transform, blind source separation, independent component analysis, convolutive mixtures

TABLE DES MATIÈRES

	Page
INTRODUCTION.....	1
CHAPITRE 1 LA SÉPARATION DE SOURCES.....	6
1.1 Introduction	6
1.1.1 Le cas aveugle	6
1.1.2 Le cas « semi-aveugle »	7
1.1.3 Historique	7
1.1.4 Les applications.....	8
1.2 Propriétés des signaux sources.....	9
1.2.1 La parole.....	10
1.2.2 Ergodisme et stationnarité.....	10
1.3 Les modèles de mélanges et de séparateurs	12
1.3.1 Le mélange instantané.....	14
1.3.2 Le mélange convolutif.....	15
1.3.3 Les systèmes de séparation	18
1.4 Les approches statistiques	20
1.4.1 L'analyse en composantes principales (ACP).....	20
1.4.2 L'analyse en composantes indépendantes (ACI)	21
1.5 Algorithmes de séparation.....	22
1.5.1 Méthode basée sur un algorithme à point fixe	23
1.5.2 Méthode basée sur une mise à jour en série.....	25
1.5.3 Méthode basée sur l'élimination de la corrélation croisée	30
1.6 Conclusion.....	35
CHAPITRE 2 LA TRANSFORMÉE EN ONDELETTES.....	37
2.1 Introduction	37
2.2 La transformée en ondelette et la transformée de Fourier à fenêtre glissante.....	37
2.3 La transformée de Fourier à temps discret et la transformée de Fourier discrète	39
2.4 La transformée de Fourier à fenêtre glissante (TFFG).....	40
2.5 La transformée de Fourier à fenêtre glissante, interprétation en banque de filtres	42
2.6 Fenêtre et Incertitude.....	43
2.7 La transformée en ondelettes.....	45
2.7.1 L'ondelette d'analyse	45
2.7.2 La transformée en ondelettes continues	47
2.8 L'analyse multirésolution.....	48
2.8.1 La banque de filtres en quadrature (deux canaux)	50
2.9 La transformée en ondelettes discrètes.....	53
2.10 La transformée en paquets d'ondelettes.....	58
2.10.1 Choix du meilleur arbre de décomposition	61
2.11 Applications de la transformée en ondelettes.....	62
2.12 Conclusion.....	63

CHAPITRE 3 LA SÉPARATION DE SOURCES À L'AIDE DE LA TRANSFORMÉE EN ONDELETTES CONTINUES.....	64
3.1 Introduction	64
3.2 Structure du séparateur.....	65
3.3 La séparation de sources avec la transformée en ondelettes continues.....	67
3.3.1 Analyse avec la transformée en ondelettes continues	69
3.3.2 Choix des échelles.....	70
3.3.3 Bloc de séparation en fréquence.....	72
3.3.4 Synthèse avec la transformée en ondelettes continues inverse	78
3.3.5 Choix de la non-linéarité.....	79
3.4 La séparation de sources avec la transformée en paquets d'ondelettes.....	79
CHAPITRE 4 SIMULATIONS	81
4.1 Introduction	81
4.2 Signaux sources.....	81
4.3 Mélanges	81
4.3.1 Mélange instantané.....	82
4.3.2 Mélanges convolutifs	82
4.4 Évaluation des performances	85
4.5 Paramètres de simulations.....	86
4.6 Résultats	88
4.7 Discussion des résultats.....	110
CONCLUSION	112
TRAVAIL FUTUR	114
BIBLIOGRAPHIE	115

LISTE DES TABLEAUX

	Page
Tableau 4.1	SS avec la TOC, mélange instantané, loi fréquences uniformes..... 89
Tableau 4.2	SS avec la TOC, mélange instantané, loi linéaire 90
Tableau 4.3	SS avec la TOC, mélange instantané, loi logarithmique..... 92
Tableau 4.4	SS avec la TOC, mélange instantané, loi mu 93
Tableau 4.5	SS avec la TOC, mélange convolutif, loi fréquences uniformes 95
Tableau 4.6	SS avec la TOC, mélange convolutif, loi linéaire 96
Tableau 4.7	SS avec la TOC, mélange convolutif, loi logarithmique..... 97
Tableau 4.8	SS avec la TOC, mélange convolutif, loi mu 98
Tableau 4.9	SS avec la TPO, mélange instantané, analyse niveau 1 99
Tableau 4.10	SS avec la TPO, mélange instantané, analyse niveau 2 100
Tableau 4.11	SS avec la TPO, mélange instantané, analyse niveau 3 101
Tableau 4.12	SS avec la TPO, mélange convolutif, analyse niveau 1 102
Tableau 4.13	SS avec la TPO, mélange convolutif, analyse niveau 2..... 103
Tableau 4.14	SS avec la TPO, mélange convolutif, analyse niveau 3 105
Tableau 4.15	SS en fréquence, mélange instantané 106
Tableau 4.16	SS en fréquence, mélange instantané 107
Tableau 4.17	Mesures de performances optimales, mélange instantané 108
Tableau 4.18	Mesures de performances optimales, mélange convolutif 109

LISTE DES FIGURES

	Page
Figure 1.1	Matrice de mélange et de séparation. 13
Figure 1.2	Séparateur avec une structure feedforward. 13
Figure 1.3	Séparateur avec une structure feedback. 14
Figure 1.4	Système de séparation direct. 19
Figure 1.5	Système de séparation en boucle. 19
Figure 1.6	Mise à jour en série d'une matrice. 25
Figure 1.7	Séparation sur deux étages. 27
Figure 2.1	Représentation temps fréquence de la TFFG. 38
Figure 2.2	Représentation temps fréquence de la TO. 39
Figure 2.3	Représentation temporelle de la TFFG. 41
Figure 2.4	Interprétation en banque de filtres de la TFFG. 42
Figure 2.5	Filtres utilisés dans l'interprétation en banque de filtres de la TFFG. ... 43
Figure 2.6	Fenêtres couramment utilisées. 44
Figure 2.7	Ondelettes couramment utilisées. 46
Figure 2.8	Décomposition du spectre par une analyse multirésolution. 50
Figure 2.9	Banque de filtres MEQ deux canaux. 51
Figure 2.10	Réponses en fréquences banque de filtres en quadrature deux canaux. 51
Figure 2.11	Banque de filtres d'analyse en octave de 3 niveaux. 54
Figure 2.12	Système équivalent 4 canaux de la figure 2.11. 55
Figure 2.13	Banque de filtres de synthèse en octave de 3 niveaux. 56
Figure 2.14	Système équivalent 4 canaux de la figure 2.13. 57
Figure 2.15	Analyse en paquet d'ondelettes. 59

Figure 2.16	Représentation équivalente de l'analyse et de la synthèse avec la TPO pour les figures 2.15 et 2.16.	60
Figure 2.17	Synthèse en paquet d'ondelettes.	61
Figure 2.18	Arbre de décomposition 2 niveaux.....	62
Figure 3.1	Algorithme de SS en fréquence.....	66
Figure 3.2	Algorithme de SS avec la TOC et la TFFG.	67
Figure 3.3	Représentation détaillée de la SS en fréquence.....	68
Figure 3.4	Représentation détaillée de la SS avec la TOC.	68
Figure 3.5	Lois de sélection des échelles.....	70
Figure 4.1	Configuration espace virtuel, vue de dessus.	83
Figure 4.2	Configuration espace virtuel, vue isométrique.....	83
Figure 4.3	Fonctions de transfert de la matrice de mélange.	84
Figure 4.4	Zéros des fonctions de transfert de la matrice de mélange.	85
Figure 4.5	Signaux sources, observés et reconstruits pour l'algorithme de SS avec la TOC utilisant les paramètres optimaux du tableau 4.2.	110

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

ACI	Analyse en composantes indépendantes
ACP	Analyse en composantes principales
EEG	Électroencéphalogramme
ECG	Électrocardiogramme
MEG	Magnétoencéphalogramme
RIF	Réponse impulsionnelle finie
RII	Réponse impulsionnelle infinie
SAS	Séparation aveugle de sources
SS	Séparation de sources
TO	Transformée en ondelettes
TOC	Transformée en ondelettes continues
TOCI	Transformée en ondelettes continues inverse
TOD	Transformée en ondelettes discrètes
TODI	Transformée en ondelettes discrètes inverse
TPO	Transformée en paquets d'ondelettes
TPOI	Transformée en paquets d'ondelettes inverse
TF	Transformée de Fourier
TFI	Transformée de Fourier inverse
TFTD	Transformée de Fourier à temps discret
TFD	Transformée de Fourier discrète
TFFG	Transformée de Fourier à fenêtre glissante
TFFGI	Transformée de Fourier à fenêtre glissante discrète inverse

LISTE DES SYMBOLES ET UNITÉS DE MESURE

a	Dilatation ou échelle
a_{ji}	Élément de la matrice de mélange
a_{lin}	Échelles, loi linéaire des échelles
a_{log}	Échelles, loi logarithmique des échelles
a_{max}	Échelle maximale
a_{min}	Échelle minimale
a_{mu}	Échelles, loi mu des échelles
a_{uni}	Échelles, loi uniforme des fréquences
A	Matrice de mélange
b	Translation
\bar{b}	Vecteur des signaux blanchis
B	Matrice de blanchissage
C	Microphone
D_l	Matrice de dilatation
D	Distorsion
$E(\cdot)$	Expectation mathématique
f_c	Fréquence centrale de l'ondelette
f_{min}	Pseudo fréquence minimale
f_{max}	Pseudo fréquence maximale
$g(\cdot)$	Non-linéarité

$H(\cdot)$	Fonction de mise à jour
i	Indice itératif
I	Matrice identité
k	Indice des échelles
k_n	Paramètre de la longueur de la TF
k_s	Paramètre de déplacement de la fenêtre
K	Nombre d'échelles
L	Niveau d'analyse avec la TPO
M	Paramètre d'évaluation de l'opérateur ε
n	Indice de temps discret
n_m	Nombre de mélanges
n_s	Nombre de sources
N	Nombre de points de la TF
N_s	Position de la fenêtre
ΔN	Déplacement de la fenêtre
\bar{p}	Coefficients de la TOC avec élimination de la permutation
P	Matrice de permutation
Q	Matrice diagonale des valeurs propres
\bar{r}	Vecteur des coefficients de la TOC reconstruits
R	Matrice de covariance
\bar{s}	Vecteur des sources
S	Qualité

t	Temps continu
U	Matrice de rotation orthogonale
\hat{U}_ω	TFFG du vecteur des coefficients de la TOC séparés
μ	Pas itératif
μ_a	Paramètre de forme, loi mu
μ_b	Paramètre de forme, loi logarithmique
V	Matrice des vecteurs propres
\hat{V}	TFFG du vecteur des coefficients de la TOC avec élimination de la dilatation
w_{ij}	Élément de la matrice de séparation
w_f	Fenêtre
W	Matrice de séparation
\bar{x}	Vecteur des mélanges
\bar{y}	Vecteur des sources reconstruites
\bar{z}	Vecteur des coefficients de la TOC
\hat{Z}	TFFG du vecteur des coefficients de la TOC
ψ	Fonction d'ondelette
σ_ω	Vecteur de permutation des échelles
σ_ω	Vecteur de permutation des fréquences
ε	Opérateur pour la résolution de la permutation
ω	Fréquence

ϕ	Fonction de coût
∇	Gradient
α	Facteur d'échelle

INTRODUCTION

Mise en situation

La séparation aveugle de sources (SAS) est un sujet de recherche récent qui a fait son apparition dans les années 1990. Son objectif est d'extraire les signaux sources à l'origine de signaux observés. Il existe beaucoup de situations où des signaux formés de mélanges sont observables. Cette constatation permet d'exposer l'étendue du domaine d'application de la SAS.

Chaque mélange est unique et nécessite une approche particulière pour en extraire les sources. Le problème est abordé différemment dépendamment de la situation. Les principaux types de mélanges sont les mélanges instantanés [1, 2] et les mélanges convolutifs [1-3]. D'autres mélanges comme le mélange non-linéaire [4] existent mais ils dépassent le cadre de ce travail. Plusieurs exemples sont présents dans la littérature [3, 5-18]. Le système de mélange possède aussi des caractéristiques temporelles dans le sens où il peut être variant ou invariant dans le temps. Clairement, le système de séparation de sources a pour objectif d'éliminer l'effet introduit par le système de mélange afin d'obtenir les sources à l'origine des mélanges.

Le nombre de signaux sources et de signaux observables forme une autre branche dans le domaine. Trois cas distincts en découlent: déterminé, sur-déterminé et sous-déterminé. Respectivement ces cas sont définis par un nombre de mélanges observés identique, supérieur ou inférieur au nombre de sources.

Objectifs

L'objectif de ce travail est de développer un algorithme de séparation de sources (SS) basé sur la transformée en ondelette continue. Ce dernier doit être en mesure d'effectuer la séparation d'un mélange convolutif dont les sources sont constituées par un locuteur et une locutrice. Les sources ont donc des distributions statistiques laplaciennes et sont considérées comme étant indépendantes. Les paramètres doivent être ajustés afin d'obtenir des performances de séparation optimales.

La dimension du problème choisi est carré dans le sens où le nombre de sources est identique au nombre de mélanges. Le cas déterminé est considéré comme étant le cas classique dans la littérature.

Démarche

Pour être en mesure de séparer le type de mélange convolutif, le système de séparation ou le séparateur doit être en mesure d'inverser la matrice à l'origine du système de mélange. Le modèle utilisé pour le système de mélange est basé sur les filtres à réponses impulsionnelles finies (RIF). Ce type de filtre permet de modéliser l'environnement acoustique à la base du mélange convolutif. L'algorithme doit donc éliminer les effets introduits par un système de mélange constitué de filtres à RIF causaux afin de reconstruire les sources d'origines.

L'algorithme développé effectue un prétraitement avec la transformée en ondelette continue (TOC) avant d'effectuer la séparation dans le domaine fréquentiel. Une transformée de fourrier à fenêtre glissante (TFFG) est appliquée aux coefficients obtenus pour chaque échelle de la TOC. La séparation a lieu en fréquence où chaque bande de fréquence est séparée pour chaque échelle de la TOC. Une permutation et une dilatation doit être éliminée

lors d'une séparation dans le domaine fréquentiel. La permutation agit sur l'appartenance des bandes de fréquence et la dilatation sur l'amplitude des celles-ci.

La permutation provient de l'incapacité du processus de séparation à déterminer l'appartenance des bandes de fréquences après séparation. Les bandes de fréquences sont séparées mais leur appartenance à une source ou à une autre est inconnue. Le problème de permutation doit être résolu indépendamment du processus de séparation. La dilatation provient de l'incapacité du processus de séparation à déterminer l'amplitude des bandes de fréquence. Cette dilatation doit être déterminée indépendamment du processus de séparation, comme dans le cas de la permutation.

Après correction de ces deux problèmes les coefficients séparés pour chaque échelle sont obtenus avec la transformée de fourrier à fenêtre glissante inverse (TFFGI). Le même problème de permutation rencontré en fréquence apparaît entre les échelles. Ce problème doit être résolu avant l'utilisation de la transformée en ondelette inverse (TOCI) pour l'obtention des signaux sources reconstruits.

Les échelles de la transformée en ondelettes sont choisies pour couvrir toute la bande passante utilisée en téléphonie. Les échelles à l'intérieur de cette bande sont sélectionnées suivant quatre lois de répartition : linéaire, logarithmique, mu et fréquences uniformes.

Dans un premier temps l'algorithme développé est appliqué au mélange instantané afin de vérifier la validité des paramètres choisis. Cette première étape permet de vérifier la convergence de l'algorithme considérant des sources avec distributions laplaciennes. Les performances de séparation sont évaluées suivant 2 paramètres : la qualité de la séparation et le niveau de distorsion. Ainsi après avoir assuré son bon fonctionnement avec le mélange instantané l'algorithme est appliqué au mélange convolutif.

L'algorithme développé est comparé au même algorithme utilisant la transformée en paquet d'ondelette (TPO) et au même algorithme sans aucun prétraitement (fréquence seulement).

Les mélanges utilisés sont synthétiques dans les deux cas. Les fonctions de transfert permettant la construction des mélanges convolutifs sont obtenues avec l'utilisation d'une boîte à outils MATLAB. Ainsi une fonction de transfert peut-être obtenue entre chaque source et chaque capteur.

Plan du mémoire

Le mémoire est constitué de quatre chapitres.

Le chapitre 1 expose les éléments essentiels entourant la séparation de sources. Il discute des concepts, des signaux sources, des types de mélanges, de séparateurs, des approches et de quelques algorithmes de séparation. Des éléments important utilisés dans l'algorithme développé proviennent des deux derniers algorithmes de séparation présentés dans ce chapitre.

Le chapitre 2 présente la théorie associée aux transformées employées dans l'ensemble du document. Les éléments les plus importants sont la TFFG, la TPO et la TOC. Quelques applications reliées à la transformée en ondelettes sont citées à la fin de ce chapitre.

Le chapitre 3 révèle les détails de la méthode de séparation de sources développée, basée sur l'utilisation de la TOC. La structure du séparateur et son fonctionnement sont présentés.

Le chapitre 4 traite des simulations effectuées en relation avec le cadre de travail choisi. Il inclut les détails concernant la procédure de simulation. Les tableaux des résultats obtenus sont présents, accompagnés d'observations générales sur les performances et le comportement de l'algorithme.

Finalement, la conclusion présente les hypothèses et observations concernant les performances de l'algorithme développé. Elle se termine par une ouverture sur les travaux futurs possibles.

CHAPITRE 1

LA SÉPARATION DE SOURCES

1.1 Introduction

L'objectif de ce chapitre est de présenter les éléments essentiels intervenant dans la SS.

La SS est une technique de traitement numérique des signaux permettant d'obtenir les composantes élémentaires (sources) à partir d'une combinaison linéaire de ces dernières telles que reçues par les capteurs utilisés [2]. L'exemple classique est l'effet cocktail où un bruit ambiant et plusieurs conversations sont présentes simultanément. La capacité de l'oreille humaine à se concentrer sur la compréhension d'une discussion particulière et d'ignorer les autres représente bien le processus de séparation. L'objectif de ce processus est l'isolation d'une source de l'ensemble des sources présentes. Notre cerveau, avec l'aide de notre système auditif est en mesure d'effectuer ce travail complexe. Le niveau de détection accompli par le système binaural peut à lui seul dans des conditions parfaites atteindre un niveau de 25 dB [19]. Cette capacité nous permet entre autre d'entretenir une conversation dans un milieu où le bruit ambiant est relativement élevé.

1.1.1 Le cas aveugle

La SAS est un cas spécifique de la SS où aucune ou peu de caractéristiques sont connues. Un bon exemple est celui où l'analyste recherche les sources significatives mais inconnues à l'origine d'un mélange. Il ne connaît rien des signaux sources omis leur existence. Un exemple pratique est l'interprétation des électroencéphalogrammes (EEG) et des magnétoencéphalogrammes (MEG). Les données recueillies sont le résultat d'un mélange de sources associé à certaines régions du cerveau. La séparation permet l'extraction des sources significatives utiles au spécialiste lors de son diagnostic. Bien entendu, l'algorithme doit être développé en collaboration avec ce dernier pour que les résultats obtenus soient utiles.

1.1.2 Le cas « semi-aveugle »

Dans certaines situations, il est possible d'émettre des hypothèses portant sur le type de sources que l'on désire reconstruire ou les caractéristiques des canaux de transmission présents. Ces éléments permettent à l'algorithme développé de converger vers une solution existante partiellement connue.

Les hypothèses au niveau des sources portent généralement sur leurs distributions. Par exemple, il est possible d'affirmer qu'une certaine source est gaussienne ou laplacienne. Au niveau du canal de transmission, les hypothèses portent généralement sur les fonctions de transfert. Ces canaux peuvent être modélisés par des filtres à réponses impulsionnelles finies (RIF) ou à réponses impulsionnelles infinies (RII). Ces hypothèses limitent l'étendu du problème transformant la séparation aveugle, où rien n'est en principe connu, en séparation « semi-aveugle ».

1.1.3 Historique

Notre capacité à porter notre attention sur un stimulus auditif particulier est un bon exemple de SS pouvant être cité. En effet nous sommes en mesure de nous concentrer sur une conversation, un son particulier en faisant abstraction des autres que nous jugeons inutiles dans une situation donnée. Les moyens que prend notre cerveau pour nous permettre de séparer différentes sources reçues sont en grande partie inconnus. Des tests ont été effectués pour tenter d'identifier certaines facettes de ce processus mais il demeure en grande partie incompris. Les premiers travaux sur le sujet ont été des expériences de perception basées sur des tests particuliers [20]. Les résultats ont démontré que notre habilité à séparer les locuteurs sont basés sur des caractéristiques telles que : le sexe du locuteur, la direction du son, sa fréquence et son débit. Arons [19] propose l'implication de facteurs autres que ceux sensoriels dans le choix de l'écoute d'un message. L'information contenue dans le message a un effet sur la sélection. En d'autres termes la sélection s'appuie sur le contenu plutôt que sur le degré de stimulation. Cette idée de sélection ou de séparation, basée sur l'information, est utilisée en SS.

Un des premiers travaux portant sur la SS et fondé sur le contenu en information est dû à Jutten et Herault [21]. Il est basé sur l'utilisation d'un réseau de neurones et de non-linéarités, et fait usage de l'analyse en composantes indépendantes (ACI). Par la suite, plusieurs algorithmes ont fait leur apparition dont certains se sont démarqués par leur efficacité. Voici quelques une de ces méthodes :

- Diagonalisation simultanée de plusieurs matrices de cumulants d'ordre 4, *Joint Approximate Diagonalisation of Eigenmatrices* (JADE) [22].
- Utilisation de l'ACI, mise à jour dont les performances sont indépendantes de la matrice de mélange, *Equivariant Adaptive Separation via Independence* (EASI) [23].
- Méthode basée sur l'approximation des distributions statistiques, algorithme de mise à jour à point fixe, *Fast Independent Component Analysis* (FAST-ICA) [24].
- Maximisation de l'information pour chaque sortie du système de séparation, *Information maximisation* (INFOMAX) [25].
- Utilisation de l'entropie comme critère de séparation, la minimisation de l'entropie correspond à une séparation optimale, *Direct Minimum Entropy Deconvolution* (DMED) [12].

Toutes ces méthodes sont basées sur les propriétés statistiques des signaux sources, le véritable point de départ des algorithmes de séparation.

1.1.4 Les applications

La SS peut être utilisée pour des fins purement analytiques ou pour extraire les signaux sources à l'origine d'un mélange. La capacité d'extraction offerte par la SS permet l'application de cet algorithme à l'ensemble des données où des signaux sources sont à la base de leur création. La séparation appliquée à toutes ces formes de mélanges, rend le domaine d'application très vaste. La SS permet de voir l'analyse de données sous un angle différent.

Voici quelques exemples d'applications :

- EEG : Élimination des artéfacts dû aux mouvements oculaires [18], élimination des artéfacts pour l'étude de l'activité cérébrale relié à des tâches mentales [17].
- MEG : Extraction de caractéristiques pour faciliter l'interprétation et la présentation [16], élimination des artéfacts [8].
- ECG du fœtus : Élimination du bruit introduit par la mère [5, 11].
- Finances : Extraction des sources significatives de revenus (facteurs cachés) [14, 24].
- Reconnaissance vocale : Amélioration des performances des systèmes de reconnaissance du locuteur par élimination des sources nuisibles [7].
- Acoustique : Obtention des sources à l'origine d'un mélange, par exemple locuteur et musique ou locuteur et locutrice [1].
- Imagerie : Élimination des réflexions [6], reconstruction d'images, analyse de documents [15].

1.2 Propriétés des signaux sources

Les signaux peuvent être divisés en deux grandes classes : les signaux déterministes et les signaux aléatoires. La première classe regroupe les signaux générés par des fonctions dont tous les paramètres sont connus. La valeur du signal est connue pour toutes les valeurs du temps. Par exemple le signal produit par la fonction sinus est déterministe. La deuxième classe regroupe les signaux dont leurs valeurs futures ne peuvent être prédites. Les valeurs de ces signaux sont aléatoires. La théorie sur les signaux aléatoires permet de définir des paramètres entourant leur comportement [23].

La parole, les signaux vidéo et audio sont des signaux aléatoires. Leurs caractéristiques sont définies par un ensemble de paramètres statistiques. Ces paramètres sont nécessaires au développement mathématique des algorithmes de SS.

1.2.1 La parole

La parole peut-être reconstruite sur une courte période par un modèle de l'appareil vocal excité par un signal glottal [3]. Cette modélisation donne naissance à deux types d'excitations, la partie voisée et la partie non-voisée. Pour la partie voisée, l'excitation est une série d'impulsions de même période que celle de la fréquence fondamentale et la partie non-voisée une excitation similaire au bruit. Les parties voisées et non-voisées couvrent respectivement et généralement les bandes de fréquences 0.1-3.5 kHz et 0.1-7 kHz [3].

Le signal de la parole est stationnaire lorsque considéré sur une courte période de temps de l'ordre de 10 millisecondes [26] et non-stationnaire lorsque considéré sur de longues périodes de l'ordre de 100 millisecondes. La non-stationnarité résulte de la variation du type d'excitation glottal et de la structure de l'appareil vocal.

1.2.2 Ergodisme et stationnarité

La SS repose sur les caractéristiques probabilistes de signaux sources et les mathématiques qui entourent les algorithmes de séparation s'appuient sur les statistiques d'ordre 2 ou d'ordre supérieur.

La moyenne et l'autocorrélation sont respectivement définies par [27]

$$u(n) = E\{x(n)\} \quad (1.1)$$

$$r(n, n-p) = E\{x(n)x^*(n-p)\} \text{ pour } p = 0, \pm 1, \pm 2, \dots \quad (1.2)$$

où * représente la conjuguée complexe.

La stationnarité au sens large implique que la moyenne et l'autocorrélation sont indépendantes du temps où elles sont évaluées. Cette définition est exprimée par [27]

$$u(n) = u, \quad (1.3)$$

$$r(n, n-p) = r(p), \quad (1.4)$$

La moyenne est constante peu importe la position dans la séquence. L'autocorrélation est indépendante de la position n , elle dépendant uniquement du délai p . La stationnarité au sens strict est très restrictive et implique l'indépendance de la distribution par rapport au temps.

La moyenne d'ensemble utilise un grand nombre de processus identiques pour évaluer la moyenne. Dans le cas où le nombre de processus est insuffisant, il est souvent possible d'observer un seul processus pendant une longue période de temps. Cette affirmation constitue la base du remplacement des moyennes d'ensemble par les moyennes temporelles. Ainsi pour un système stationnaire et ergodique les moyennes d'ensemble sont égales aux moyennes temporelles. La mesure des propriétés statistiques devient donc possible à partir d'un seul système avec un nombre suffisant d'échantillons.

Pour la séquence $x(n)$ avec $n=1,2,\dots,N_e$ la moyenne temporelle et l'autocorrélation temporelle sont [27]

$$\hat{m}_t = \frac{1}{N_e + 1} \sum_{n=0}^{N_e} x(n). \quad (1.5)$$

$$\hat{r}_t(p) = \frac{1}{N_e + 1} \sum_{n=0}^{N_e} x(n)x^*(n-p). \quad (1.6)$$

Un processus stationnaire aléatoire est ergodique au niveau de la moyenne si la moyenne d'ensemble est égale à la moyenne temporelle pour la limite $N_c \rightarrow \infty$,

$$m_r = \lim_{N_c \rightarrow \infty} \hat{m}_r. \quad (1.7)$$

Un procédé stationnaire aléatoire est ergodique au niveau de la corrélation si l'autocorrélation est égale à la corrélation temporelle pour la limite $N_c \rightarrow \infty$, comme [27]

$$r_r(p) = \lim_{N_c \rightarrow \infty} \hat{r}_r(p) \quad (1.8)$$

Il faut noter que l'ergodicité implique la stationnarité.

Les algorithmes d'ordre 2 appliqués à des signaux stationnaires peuvent quand même être appliqués à des signaux stationnaires à court terme. Les statistiques sont évaluées sur la période de stationnarité des observations [28, 29].

1.3 Les modèles de mélanges et de séparateurs

Définissons les variables suivantes : n_s le nombre de sources et n_m le nombre de mélanges. Ces deux quantités définissent les dimensions du problème et génèrent trois cas possibles : le cas déterminé $n_m = n_s$, le cas sous-déterminé $n_m < n_s$ et le cas sur-déterminé $n_m > n_s$. Le cas typique $n_m = n_s = 2$ est représenté à la figure 1.2. La figure 1.1 est une représentation de plus haut niveau que la figure 1.2. La figure 1.1 expose le modèle dans son ensemble tandis que la figure 1.2 expose les éléments de la matrice de mélange A et de séparation W .

La matrice de mélange A est à l'origine des signaux observés $\bar{x}(n)$. Elle est le modèle du canal de transmission existant entre les signaux sources $\bar{s}(n)$ et les signaux observés $\bar{x}(n)$.

La matrice de séparation W permet d'obtenir les signaux sources reconstruits $\bar{y}(n)$ à partir des signaux observés $\bar{x}(n)$.

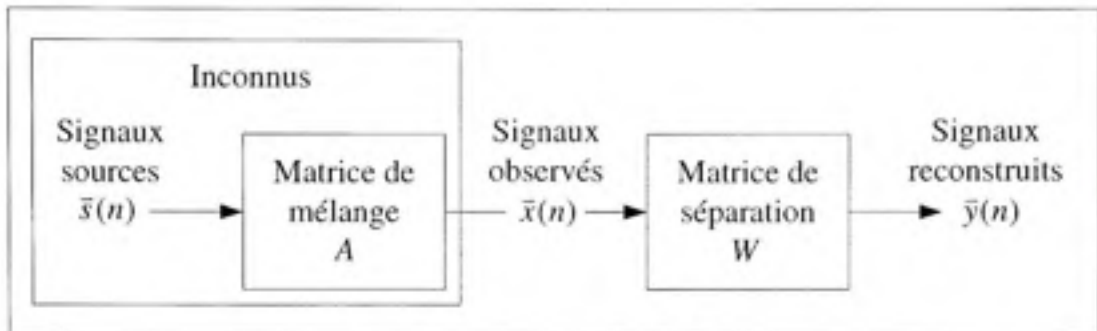


Figure 1.1 *Matrice de mélange et de séparation.*

Deux structures sont présentées dans la littérature [1] pour la réalisation du séparateur : la structure *feedforward* (figure 1.2) et la structure *feedback* (figure 1.3). La structure *feedback* est moins stable que la structure *feedforward*. Ceci est dû à la rétroaction introduite par les éléments croisés. La structure *feedforward* possède l'avantage d'être facilement implantable dans le domaine fréquentiel en employant la méthode *overlap and save* [3].

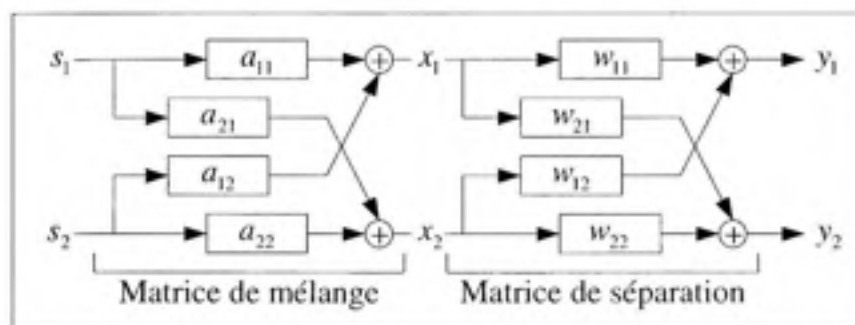


Figure 1.2 *Séparateur avec une structure feedforward.*

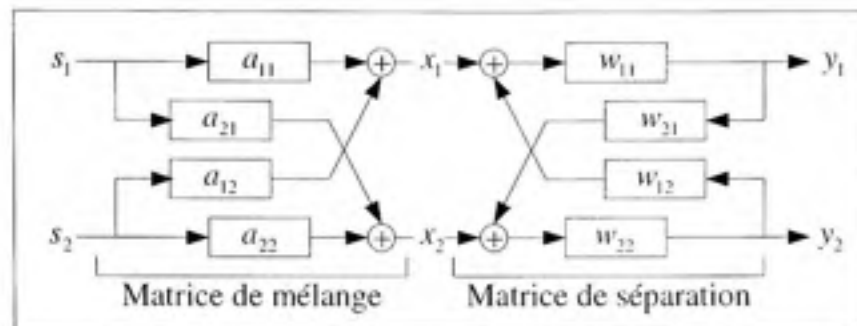


Figure 1.3 Séparateur avec une structure feedback.

Les figures 1.2 et 1.3 représentent des structures de séparateurs qui peuvent être appliqués autant au niveau du mélange instantané que convolutif. Ce sont seulement les coefficients des matrices A et W qui changent de formes comme exposé dans la section 1.3.2.

1.3.1 Le mélange instantané

La majorité des algorithmes présents dans la littérature ont été développés pour traiter le cas des mélanges instantanés, ce dernier étant le modèle le plus simple de mélange [30]. Ces mélanges sont caractérisés par des délais inexistant, négligeables ou identiques entre les sources et les capteurs. La matrice de mélange A est constituée de coefficients a_{ji} représentant les poids des mélanges. Les sources $s_i(n)$ et les poids de la matrice A sont des inconnus. Dans la réalité, A peut être une combinaison des propriétés suivantes : linéaire, non-linéaire et variant, invariant dans le temps. Le cas généralement considéré est celui du système linéaire et invariant.

Les mélanges sont obtenus par

$$x_j(n) = \sum_i a_{ji} s_i(n) \text{ pour } i = 1, \dots, n_s, j = 1, \dots, n_m \quad (1.9)$$

et sous la forme matricielle par

$$\bar{x}(n) = A\bar{s}(n) . \quad (1.10)$$

où

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n_s} \\ a_{21} & a_{22} & \cdots & a_{2n_s} \\ \vdots & & & \vdots \\ a_{n_m 1} & a_{n_m 2} & \cdots & a_{n_m n_s} \end{bmatrix} \quad (1.11)$$

La matrice de séparation W est sélectionnée pour permettre l'inversion du système inconnu A . Pour la structure *feedforward* de la figure 1.2 les sources reconstruites sont définies par

$$y_i(n) = \sum_j w_{ij} x_j(n) \text{ pour } i = 1, \dots, n_s \quad j = 1, \dots, n_m \quad (1.12)$$

et sous la forme matricielle par

$$\bar{y}(n) = W\bar{x}(n) . \quad (1.13)$$

L'insertion de (1.10) dans (1.13) permet d'obtenir le système global défini par

$$\bar{y}(n) = W\bar{x}(n) = WA\bar{s}(n) . \quad (1.14)$$

1.3.2 Le mélange convolutif

Si l'on remplace les coefficients de A par des filtres [1-3] le modèle devient celui du mélange convolutif. Les multiplications précédentes sont remplacées par l'opérateur de

convolution. Tout ce qui a été dit précédemment s'applique toujours pour la séparation des sources. La matrice de mélange A est construite pour modéliser le canal de transmission pour lequel l'algorithme de séparation est développé.

Dans le domaine de l'acoustique, les canaux sont habituellement modélisés par des filtres à réponses impulsionnelles finie (RIF) [1, 3]. Ils permettent une somme des sources avec des amplitudes et des délais différents dans le temps. Ce modèle représente bien les réflexions multiples et l'atténuation causée par les différents matériaux. Chaque élément a_{ji} de la matrice est un filtre RIF modélisant les propriétés acoustiques de l'environnement.

Donc

$$x_j(n) = \sum_i a_{ji} * s_i(n) \text{ pour } i = 1, \dots, n_s \quad j = 1, \dots, n_m \quad (1.15)$$

où a_{ji} sont des filtres RIF.

Dans le cas où $n_m = n_s = 2$, le développement de la formule (1.11) donne

$$\begin{aligned} x_1(n) &= a_{11} * s_1(n) + a_{12} * s_2(n) \\ x_2(n) &= a_{12} * s_1(n) + a_{22} * s_2(n) \end{aligned} \quad (1.16)$$

La matrice W doit permettre l'élimination des effets introduits par le canal de transmission. Elle doit être en mesure d'égaliser l'inverse de la matrice A pour permettre une séparation efficace. Différents types de filtres peuvent être employés dépendamment du modèle du canal.

Les filtres w_{ij} peuvent être des filtres RIF ou RII. Chaque choix possède ses avantages et inconvénients. Ils peuvent être utilisés dans la structure *feedforward* ou *feedback*. Dans le cas des filtres RII, l'avantage est de pouvoir traiter l'inversion de réponses impulsionnelles

beaucoup plus longues [1]. Les réponses acoustiques possèdent généralement une longueur de plusieurs milliers d'échantillons (ou de plusieurs msec) [31]. La longueur de l'inverse dépend des propriétés de ces réponses. L'inconvénient de l'utilisation de filtres RII dans le bloc de séparation est la stabilité. Les filtres RIF sont toujours stables, ce qui n'est pas le cas des filtres RII.

Une attention particulière doit être portée à la matrice de mélange. Lorsque tous les zéros des filtres a_{ji} sont à l'intérieur du cercle unité les filtres sont à phase minimale. Ce qui implique que les filtres inverses sont stables. Les zéros des filtres a_{ji} deviennent les pôles des filtres w_{ij} . Lorsque les fonctions de transferts a_{ji} ont des zéros à l'extérieur du cercle unité leurs inverses seront des fonctions de transferts à phases non-minimales. Le système de séparation doit être conçu pour permettre l'inversion d'un système à déphasage non-minimal (séparateur non-causal). Par exemple, un système à déphasage non-minimal est présent lorsque l'écho capté par un microphone a une amplitude plus grande que le signal direct [32]. Des méthodes utilisant l'introduction d'un délai dans la boucle d'apprentissage [32] ou l'utilisation de la *backpropagation* [33] permettent le traitement des filtres non-causaux.

Dans le cas où les matrices A et W sont tous des filtres RIF et que W possède une structure *feedforward*

$$y_i(n) = \sum_j w_{ij} * x_j(n) \text{ pour } i = 1, \dots, n_s \quad j = 1, \dots, n_w \quad (1.17)$$

où w_{ij} sont des filtres RIF. Après développement pour $n_w = n_s = 2$ les signaux reconstruits sont

$$\begin{aligned} y_1(n) &= w_{11} * x_1(n) + w_{12} * x_2(n) \\ y_2(n) &= w_{21} * x_1(n) + w_{22} * x_2(n) \end{aligned} \quad (1.18)$$

Il est possible de traiter le cas du mélange par convolution dans le domaine fréquentiel en remplaçant les éléments par leur transformée de Fourier (TF) [3] et les convolutions par des multiplications. L'avantage est l'expression simplifiée des convolutions et la diminution du temps de calcul.

Dans le domaine fréquentiel les formules (1.17) et (1.18) sont respectivement définies par

$$Y_i = \sum_j W_{ij} X_j \text{ pour } i = 1, \dots, n_s, j = 1, \dots, n_m \quad (1.19)$$

et

$$\begin{aligned} Y_1 &= W_{11}X_1 + W_{12}X_2 \\ Y_2 &= W_{21}X_1 + W_{22}X_2 \end{aligned} \quad (1.20)$$

1.3.3 Les systèmes de séparation

Deux approches existent [2] pour la construction du système de séparation : l'approche directe (figure 1.4) et l'approche en boucle (figure 1.5). L'approche directe ou hors ligne détermine les paramètres du séparateur à partir des mélanges et d'un bloc annexe de traitement. Elle calcule le séparateur à partir des mélanges recueillis. Les méthodes de diagonalisation entrent dans cette catégorie de séparateur. Un exemple est l'algorithme JADE qui diagonalise la matrice de quadricovariance. Un des avantages de la méthode en bloc est de permettre l'estimation des paramètres statistiques en utilisant l'espérance mathématique $E(\cdot)$.

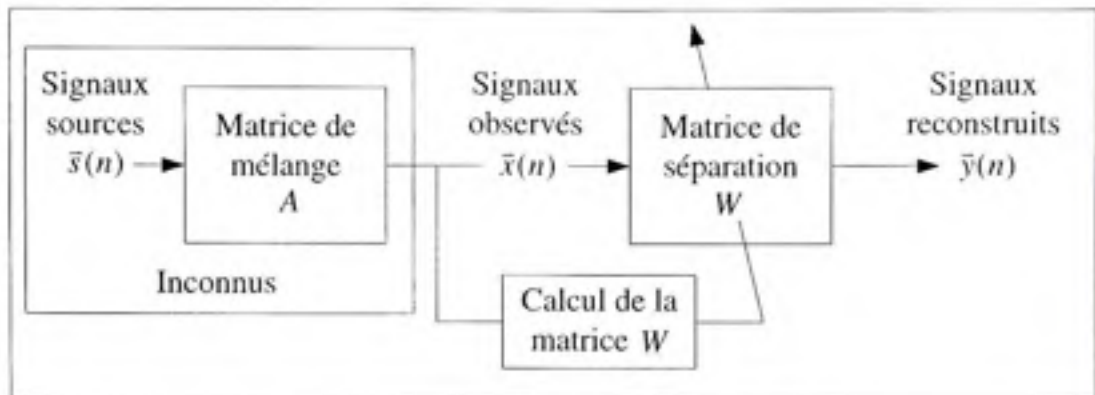


Figure 1.4 *Système de séparation direct.*

L'approche en boucle ou en ligne utilise la rétroaction pour adapter le séparateur. Les réseaux neuromimétiques font partie de cette catégorie [2]. Les deux approches peuvent être unifiées pour créer une approche bloc en ligne [34].

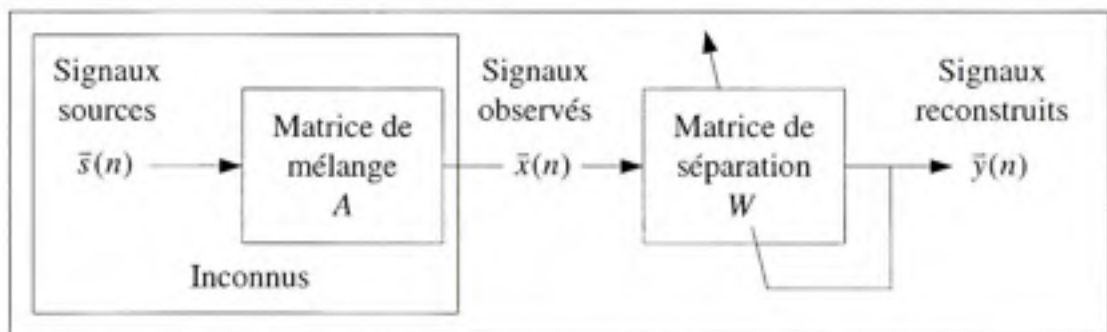


Figure 1.5 *Système de séparation en boucle.*

Il est important de noter que dans le cas du mélange convolutif la séparation est reliée au domaine de l'identification paramétrique. Elle peut-être vue comme l'identification d'un système de filtres à RIF ou à RII. La séparation est aussi un cas d'égalisation aveugle étendu à plusieurs canaux. Les techniques développées dans un domaine sont applicable à l'autre et vice versa.

1.4 Les approches statistiques

La SS basée sur les statistiques d'ordre 2 [1] repose uniquement sur la décorrélation des sources reconstruites. Ces statistiques n'exploitent pas entièrement l'hypothèse d'indépendance des sources habituellement utilisée. En effet, la décorrélation (blanchiment) ne garantit pas l'indépendance des signaux. Les statistiques d'ordre supérieur à 2 entrent en jeux pour permettre l'exploitation complète de l'indépendance. Une meilleure exploitation de celle-ci implique une amélioration de la séparation. L'ACI exploite ces statistiques ordres supérieurs complétant ainsi l'analyse en composantes principales (ACP), qui est une méthode basée sur les statistiques d'ordre 2.

1.4.1 L'analyse en composantes principales (ACP)

L'ACP est une méthode permettant d'obtenir des signaux décorrélés à la sortie d'un système [1]. Cette analyse trouve les directions de plus grandes variances dans les données et projette les données ces axes. Les nouvelles données obtenues sont orthogonales donc décorrélées.

La matrice de séparation est calculée à partir des vecteurs propres de la matrice de covariance définis par [1]

$$VQV^{-1} = E[\bar{x}\bar{x}^T] - E[\bar{x}]E[\bar{x}]^T \quad (1.21)$$

où V est une matrice des vecteur propres et Q une matrice diagonale des valeurs propres de la matrice de covariance.

La matrice de séparation W est définie par [1]

$$W = Q^{-1/2}V^T \quad (1.22)$$

L'ACP est une solution orthogonale qui ne sépare pas nécessairement les mélanges. Une opération de rotation qui préserve l'orthogonalité est nécessaire à la suite de la décorrélation.

1.4.2 L'analyse en composantes indépendantes (ACI)

L'ACI [24] est une méthode statistique permettant d'obtenir, par transformation de variables aléatoires, des variables indépendantes. Cette approche permet de définir un modèle dans lequel les variables inconnues sont indépendantes et le système de mélange est inconnu. Le système de mélange est à l'origine des variables aléatoires observées.

Le séparateur utilise une transformation inverse pour reconstruire les sources d'origine. La transformation ajuste les propriétés statistiques des sources reconstruites jusqu'à l'obtention de celles recherchées. Ces caractéristiques peuvent être : un type de distribution, des moments statistiques particuliers, un niveau d'entropie, etc.

Une des approches est basée sur l'analyse de la distribution des variables. Le signal de la parole possède une distribution laplacienne plutôt que gaussienne [35]. La plupart des valeurs sont situées autour de zéro et des extrêmes. Une mesure de l'indépendance, dans ce cas, est donc une mesure de la « non-gaussiannité ». Le théorème central limite stipule que : « la somme de variables non-gaussiennes possède une distribution plus gaussienne que celle d'origine » [9]. La distribution des mélanges est donc plus gaussienne que la distribution des sources. Une maximisation de la non-gaussianité des sources reconstruites par l'utilisation du principe du maximum de vraisemblance a pour effet de maximiser l'indépendance des signaux. L'application de ce principe permet d'évaluer la distance d'une distribution par rapport à une autre (dans ce cas gaussienne) par l'entremise de la divergence Kullback-Leibler [1] qui est un estimateur de vraisemblance.

Voici une liste des principales méthodes utilisées en ACI pour accomplir la séparation :

- Méthode basée sur les cumulants croisés et la quadricovariance
 - Extension des méthodes de diagonalisation d'ordre 2; diagonalisation d'une matrice cumulante [22].
- Méthodes basées sur des contrastes
 - La maximisation de l'entropie; maximisation de l'information individuelle des signaux à la sortie d'un système [36].
 - Minimisation de l'information mutuelle; minimisation de l'information entre les signaux à la sortie d'un système [37].
- Méthodes neuromimétiques
 - Exploitation des ordres supérieurs due à l'utilisation de non-linéarités [25].
- Méthode utilisant le maximum de vraisemblance
 - La maximisation de la non-gausiannité; maximisation de la distance par rapport à une distribution gaussienne [38].

Il est intéressant de noter que la séparation peut être associée à un problème d'optimisation. Une fonction de coût basée sur un critère d'indépendance est optimisée avec une méthode itérative comme le gradient. L'optimum de la fonction de coût est le point où les sources reconstruites répondent aux critères statistiques désirés.

1.5 Algorithmes de séparation

Comme vu au début de la section 1.4, les algorithmes de SS sont divisés en deux grandes classes : les méthodes d'ordre 2 et les méthodes d'ordres supérieurs. Dans le premier cas, une méthode connue est la diagonalisation simultanée de plusieurs matrices de corrélations à plusieurs délais [26, 39, 40]. Cette approche élimine la corrélation croisée ou la corrélation entre les différents signaux observés. Dans le second cas, les travaux dans le domaine convergent vers l'utilisation de non-linéarités qui définissent le comportement de l'algorithme. Elles ont une influence directe sur sa stabilité et sa convergence [23, 37].

Le domaine temporel n'est pas le seul domaine dans lequel il est possible d'effectuer la SS. D'autres domaines s'offrent, chacun possédant des caractéristiques particulières. Il est avantageux de choisir un domaine offrant la représentation la plus appropriée par rapport au problème posé. Ce transfert, dépendamment des circonstances, permet entre autre : d'utiliser d'autres propriétés, de diminuer la complexité du problème et d'augmenter la vitesse de traitement. En plus des transformées, la bande spectrale occupée par le signal peut-être divisée en sous bandes avant d'effectuer la séparation. Comme la transformée cette division modifie les performances de l'algorithme global. Ainsi, l'utilisation des coefficients cepstraux [41], de la transformée en ondelettes discrètes (TOD) [10, 42-47], de banques de filtres [48, 49], du spectre de Wigner-Ville [28] et de la TFFG [13, 50] apparaît dans la littérature. Au cœur de chacune de ces approches réside un algorithme de séparation basé sur l'hypothèse d'indépendance statistique des sources.

Les méthodes décrites dans les sections 1.51, 1.52 et 1.53 ont pour objectif d'exposer les éléments importants des méthodes qui y sont décrites. Certains de ces éléments sont utilisés dans les chapitres suivants.

1.5.1 Méthode basée sur un algorithme à point fixe

Cet algorithme [24] propose l'utilisation d'une nouvelle gamme de fonctions de coûts basées sur l'approximation de l'entropie différentielle. L'entropie différentielle peut être définie comme le niveau d'information véhiculé par une variable aléatoire.

L'optimisation des fonctions est opérée par un algorithme à point fixe. Cette méthode combine deux approches existantes pour atteindre son objectif: l'ACI et la poursuite des projections. L'ACI est effectuée en minimisant l'information mutuelle entre les variables transformées et les fonctions de contrastes sont évaluées du point de vue de la poursuite de projection en permettant une extraction séquentielle des sources. L'information mutuelle est une quantité qui permet de mesurer le niveau de dépendance entre deux variables. Tous ces éléments sont développés et appliqués au cas du mélange instantané.

La non-linéarité $g(\cdot)$ dans l'algorithme (1.23) dépend de l'information véhiculé par les sources recherchées. Elle est choisie pour modéliser la fonction de répartition des sources à reconstruire. La matrice de séparation du point de départ de l'algorithme, doit être initialisée. Dans la majorité des cas cette matrice initiale est choisie de façon aléatoire.

L'algorithme à point fixe est défini par [24]

$$\begin{aligned} W_{i+1} &= E[\bar{x}g(W_i^T x)] - E[g'(W_i^T x)]W_i \\ W_{i+1} &= W_{i+1} / \|W_{i+1}\| \end{aligned} \quad (1.23)$$

et la version stabilisée par [24]

$$\begin{aligned} W_{i+1} &= W_i - \mu [E\{\bar{x}g(W_i^T x)\} - \beta W_i] / [E\{g'(W_i^T x)\} - \beta] \\ W_{i+1} &= W_{i+1} / \|W_{i+1}\| \end{aligned} \quad (1.24)$$

où $\beta = E[W_i^T \bar{x}g(W_i^T x)]$, W_i la matrice de séparation actuelle, W_{i+1} est la matrice de séparation mise à jour, μ le pas et $g(\cdot)$ une fonction non-linéaire représentant la fonction de répartition des sources recherchées. Si $\mu = 1$ l'algorithme stabilisé (1.24) redevient celui de base (1.23).

Les propriétés de l'algorithme sont les suivantes :

- La convergence est cubique ce qui augmente la vitesse par rapport aux algorithmes basés sur le gradient qui possèdent une convergence linéaire.
- L'algorithme original (point fixe) ne possède pas le paramètre du pas ce qui augmente la stabilité de l'algorithme final (stabilisé).
- L'utilisation de non-linéarités permet l'obtention de sources non-gausiennes.
- Le choix de la non-linéarité permet l'optimisation.

- L'algorithme permet l'obtention des composantes indépendantes une par une (extraction séquentielle).
- L'algorithme hérite des propriétés des algorithmes basés sur les réseaux de neurones.

1.5.2 Méthode basée sur une mise à jour en série

L'algorithme EASI [23] est basé sur la mise à jour en série (figure 1.6) des éléments nécessaires à la séparation. La performance de cette approche ne dépend pas de la matrice de mélange A . Cette propriété est la plus importante.

Cette méthode effectue la SS avec un système de séparation réparti sur deux étages (figure 1.7). La matrice de séparation W est décomposée en deux matrices : une matrice de blanchissage B et une matrice de rotation U . La matrice de blanchiment est un prétraitement qui permet l'orthogonalisation des mélanges. La matrice de rotation permet de séparer les signaux en conservant la contrainte d'orthogonalité sur les signaux reconstruits. La performance uniforme (indépendante de la matrice A) est garantie par la mise à jour en série des deux matrices. La performance générale de l'algorithme est reliée au choix des non-linéarités $\bar{g}(\cdot)$, choisies en fonction des distributions des sources recherchées.

L'algorithme de mise à jour est basé sur la définition d'une fonction $H(\cdot)$ permettant la mise à jour de W avec

$$W_{i+1} = (I - \mu H_i(\bar{y}))W_i. \quad (1.25)$$

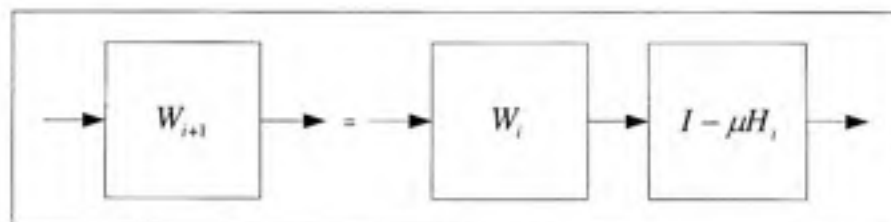


Figure 1.6 Mise à jour en série d'une matrice.

La fonction $H(\cdot)$ est obtenue en évaluant le gradient relatif de la fonction de coût. Pour cet algorithme la fonction de coût est basée sur l'utilisation des cumulants.

La fonction de coût est définie par [23]

$$\phi_c(B) \stackrel{\text{def}}{=} E[f(y)] \text{ avec } f(y) = \sum_{i=1, n_M} |y_i|^4 \quad (1.26)$$

et son gradient relatif par

$$\nabla \phi_c(B) = \nabla E f(\bar{y}) = \nabla E f(W\bar{x}) = E[f'(\bar{y})\bar{y}^T]. \quad (1.27)$$

La mise à jour de (1.25) par la fonction $H(\cdot)$ obtenue en (1.27) devient [23]

$$W_{i+1} = W_i - \mu E[f'(\bar{y})\bar{y}^T] W_i. \quad (1.28)$$

L'équation (1.28) représente la mise à jour en série de la matrice de séparation sans étape de blanchiment. C'est le résultat du calcul direct du gradient de la fonction de coût (1.26).

Comme exposé précédemment l'algorithme final est composé d'un bloc de blanchiment et d'un bloc de séparation. La première étape est le blanchiment des mélanges par la matrice B . La deuxième étape est la séparation des signaux blanchis par la matrice U . Cette matrice doit optimiser la fonction de coût (1.26) sous la contrainte de décorrélation (orthogonalité) imposée par B .

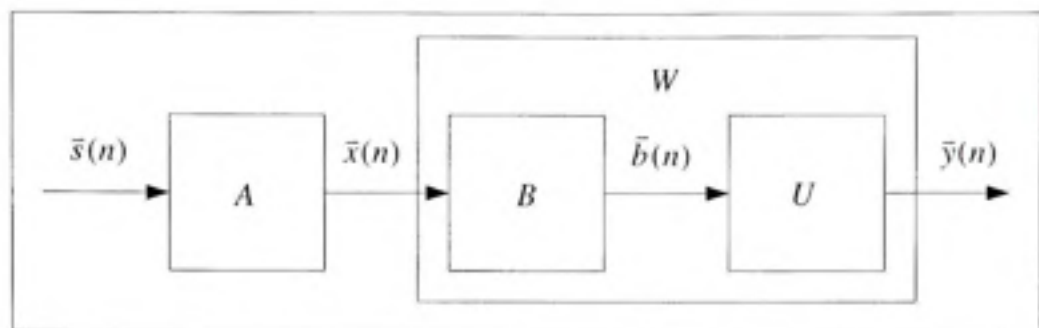


Figure 1.7 Séparation sur deux étages.

La matrice B est une matrice dont les sorties $\bar{z}(t)$ sont blanchies. Elle est définie par

$$R_z = BR_x B^T \quad (1.29)$$

où $R_x = E[\bar{x}\bar{x}^T]$ la matrice de covariance des mélanges et $R_z = E[\bar{z}\bar{z}^T]$ la matrice de covariance des mélanges blanchis. L'objectif est l'obtention d'une transformation B qui diagonalise la matrice de covariance des mélanges. La matrice B est obtenue en minimisant la distance entre R_z et la matrice identité I .

La divergence de Kullback-Leibler [1] entre deux distributions normales de moyennes nulles avec des matrices de covariances égales à R_z et I est [23]

$$KL(R_z, I) \stackrel{\text{def}}{=} \frac{1}{2} (\text{trace}(R_z) - \log \det(R_z) - n). \quad (1.30)$$

La propriété la plus importante est que $KL(R_z, I) \geq 0$ avec égalité si et si seulement si $R_z = I$. La matrice B sera une matrice de blanchiment si $KL(R_z, I) = 0$.

Elle doit minimiser

$$\phi_B(B) = KL(BR_s B^T, I). \quad (1.31)$$

Le gradient relatif de ϕ_B est défini par [23]

$$\nabla \phi_B \stackrel{def}{=} R_z - I \quad (1.32)$$

et l'algorithme de mise à jour en série de la matrice de blanchiment par [23]

$$B_{i+1} = B_i - \mu E[\bar{b}\bar{b}^T - I]B_i. \quad (1.33)$$

L'algorithme doit permettre la mise à jour de la matrice U en préservant la décorrélation imposée par la matrice B . La matrice U doit être orthogonale et doit conserver cette forme tout au long du processus d'adaptation. La mise à jour de la matrice U sous la contrainte d'orthogonalité est définie par [23]

$$U_{i+1} = U_i - \mu E[f'(\bar{y})\bar{y}^T - \bar{y}f'(\bar{y})^T]U_i. \quad (1.34)$$

L'équation (1.34) est le résultat de l'utilisation de la contrainte d'orthogonalité sur la mise à jour sans contraintes (1.28). La combinaison des deux blocs $W_{i+1} = U_{i+1}B_{i+1}$ avec l'utilisation de $U_i(I - \mu[\bar{b}\bar{b}^T - I]) = (I - \mu[\bar{y}\bar{y}^T - I])U_i$, $U_i U_i^T = I$ et $\bar{y} = U_i \bar{b}$ donne la fonction $H(\cdot)$ suivante [23]

$$H(y) = \bar{y}\bar{y}^T - I + f'(\bar{y})\bar{y}^T - \bar{y}f'(\bar{y})^T. \quad (1.35)$$

Elle comprend les opérations de blanchiment et de séparation. Le premier et le second terme représentent le blanchissage, les troisième et quatrième termes représentent la séparation. Les deux derniers termes apportent les conditions nécessaires autres que d'ordre 2 pour finaliser la séparation. L'algorithme converge lorsque les éléments hors diagonale de $H(y)$ deviennent nuls. Elle est atteinte lorsque les signaux sont blanchis et que la corrélation non-linéaire croisée est nulle $offdiag(E[\bar{y}f'(\bar{y})^T])=0$. Pour un ensemble de non-linéarités

$$g(\bar{y}) = [g_1(y_1), \dots, g_{n_g}(y_{n_g})]^T, \quad (1.36)$$

l'algorithme EASI est défini par [23]

$$W_{i+1} = W_i - u [\bar{y}\bar{y}^T - I + g(\bar{y})\bar{y}^T - \bar{y}g(\bar{y})^T] W_i, \quad (1.37)$$

et la version normalisée par [23]

$$W_{i+1} = W_i - u \left[\frac{\bar{y}\bar{y}^T - I}{1 + \mu \bar{y}^T \bar{y}} + \frac{g(\bar{y})\bar{y}^T - \bar{y}g(\bar{y})^T}{1 + \mu |\bar{y}^T g(\bar{y})|} \right] W_i, \quad (1.38)$$

où W_{i+1} est la matrice de séparation mise à jour, \bar{y} les sources reconstruites, $g(\cdot)$ une fonction non-linéaire, I la matrice identité et u le pas.

Les propriétés de l'algorithme sont les suivantes :

- La structure de mise à jour est simple.
- La performance ne dépend pas de la matrice de mélange.
- La vitesse de convergence, la stabilité et le niveau de séparation dépendent de la distribution (normalisée) des signaux sources.
- Pour chaque non-linéarité il existe un point de stationnarité.
- La version stabilisée n'impose pas de contrainte sur la matrice de séparation.

1.5.3 Méthode basée sur l'élimination de la corrélation croisée

Cette méthode [26] est basée sur l'élimination de la corrélation entre les mélanges à différent délais. Les matrices de corrélation sont calculées pour chaque délai et diagonalisées par l'utilisation d'un algorithme de diagonalisation jointe. La méthode est implantée dans le domaine fréquentiel où elle est en mesure de traiter le mélange convolutif.

La première étape est le blanchiment des mélanges. Cette transformation rend orthogonales les directions des distributions des points de ces mélanges. Cette opération ne permet pas à elle seule la séparation des sources, une rotation doit être introduite. La rotation est assurée par l'utilisation d'une matrice de transformation. L'obtention de cette matrice est basée sur la diagonalisation jointe de plusieurs matrices de corrélation des observations blanchies, obtenues à différent délais comme cité précédemment.

Pour appliquer cette méthode au mélange convolutif, le passage au domaine fréquentiel est nécessaire. L'objectif est la séparation de chacune des composantes fréquentielles de la TFFG avec la méthode exposée précédemment. L'obtention d'une matrice contenant l'opération de blanchissage et de rotation permet d'effectuer cette séparation. Avant d'effectuer la reconstruction des sources par la TFFGI deux problèmes doivent être résolus : la dilatation et la permutation.

Ces éléments ont un impact important dans le domaine fréquentiel, chaque composante est affectée ce qui cause une erreur globale lors de la reconstruction des sources. La balance spectrale est affectée par la dilatation (amplification) et l'appartenance des composantes fréquentielle à chaque source par la permutation. Après résolution de ces problèmes les sources sont reconstruites avec la TFFGI.

La dilatation et la permutation sont respectivement représentées par les matrices P et D_i . Tous les éléments de P pour chaque colonne et chaque ligne sont nuls à l'exception d'un seul prenant la valeur de 1 et D_i est une matrice diagonale.

Le système dilatation permutation PD_f est égal au système global WA comme

$$PD_f = WA. \quad (1.39)$$

Le mélange par convolution est défini par

$$\bar{x}(n) = A * \bar{s}(n) \quad (1.40)$$

et dans le domaine fréquentiel par

$$\hat{\bar{X}}(\omega) = A(\omega)\hat{\bar{S}}(\omega). \quad (1.41)$$

La première étape, l'application de la TFFG

$$\hat{\bar{X}}(\omega, N_s) = \sum_n e^{-j\omega n} \bar{x}(n) w_f(n - N_s) \quad (1.42)$$

avec

$$\omega = 0, \frac{1}{N} 2\pi, \dots, \frac{N-1}{N} 2\pi, \quad N_s = 0, \Delta N, 2\Delta N, \dots \quad (1.43)$$

où ω est la fréquence, N le nombre de point de la transformée de Fourier discrète (TFD), N_s la position de la fenêtre dans le temps, $w_f(\cdot)$ la fenêtre utilisée et ΔN_s le déplacement de la fenêtre.

La TFFGI est définie par

$$\bar{x}(n) = \frac{1}{2\pi} \frac{1}{W_f(n)} \sum_{N_s} \sum_{\omega} e^{-j\omega(n-N_s)} \hat{\bar{X}}(\omega, N_s), \quad (1.44)$$

où

$$W_f(n) = \sum_{N_s} w_f(n - N_s). \quad (1.45)$$

Après transformations de chaque élément de (1.41) le système de mélange devient

$$\hat{\bar{X}}(\omega, N_s) = \hat{A}(\omega) \hat{\bar{S}}(\omega, N_s), \quad (1.46)$$

où $\hat{A}(\omega)$ est la TF de A , et $\hat{\bar{S}}(\omega, N_s)$ la TFFG des sources $\bar{s}(n)$.

Si la fréquence est fixée la notation devient

$$\hat{\bar{X}}_{\omega}(N_s) = \hat{\bar{X}}(\omega, N_s). \quad (1.47)$$

Après séparation, les composantes fréquentielles sont mutuellement indépendantes pour chaque fréquence ω . Les composantes fréquentielles séparées sont

$$\hat{\bar{U}}_{\omega}(N_s) = W(\omega) \hat{\bar{X}}_{\omega}(N_s) \quad (1.48)$$

où $W(\omega)$ est la matrice de séparation. La dilatation est résolue par l'application de l'inverse de la matrice de séparation à chacune des composantes fréquentielles séparées.

Chaque spectrogramme est divisé en deux et défini par

$$\tilde{\bar{V}}_{\omega}(N_s; i) = W(\omega)^{-1} \begin{pmatrix} 0 \\ \vdots \\ \tilde{U}_{i,\omega}(N_s) \\ \vdots \\ 0 \end{pmatrix}. \quad (1.49)$$

La permutation des composantes spectrales est résolue par l'utilisation de la propriété de non-stationnarité des signaux sources. Aucune corrélation n'existe entre les différentes fréquences du spectrogramme lorsque les signaux sont stationnaires. Cependant, la non-stationnarité des signaux sources permet l'exploitation d'une propriété importante, si les bandes proviennent de la même source elles sont sous l'influence de la même modulation en amplitude. Par corrélation des enveloppes pour chaque fréquence, il devient possible de retrouver l'appartenance des bandes à une ou l'autre des sources reconstruites [26].

L'opérateur permettant d'obtenir les enveloppes est défini par

$$\varepsilon \tilde{\bar{V}}_{\omega}(t_s; i) = \frac{1}{2M} \sum_{N'_s=N_s-M}^{N_s+M} \sum_{j=1}^n |\tilde{V}_{j,\omega}(N'_s; i)|, \quad (1.50)$$

où M est une constante positive et $\tilde{V}_{j,\omega}(N_s; i)$ signifie l'élément j du vecteur $\tilde{\bar{V}}_{\omega}(N_s; i)$. Le produit scalaire de deux opérateurs est défini par

$$\varepsilon \tilde{\bar{V}}_{\omega}(i) \cdot \varepsilon \tilde{\bar{V}}_{\omega'}(j) = \sum_{N_s} \varepsilon \tilde{\bar{V}}_{\omega}(N_s; i) \varepsilon \tilde{\bar{V}}_{\omega'}(N_s; j) \quad (1.51)$$

et sa norme par

$$\|\varepsilon \tilde{\bar{V}}_{\omega}(i)\| = \sqrt{\varepsilon \tilde{\bar{V}}_{\omega}(i) \cdot \varepsilon \tilde{\bar{V}}_{\omega}(i)}. \quad (1.52)$$

Voici la méthode de résolution de la permutation étape par étape [26] :

1. Trier ω en ordre de grandeur de la corrélation entre les éléments indépendants dans ω . Cette étape est accomplie en ordonnant les fréquences comme ci-dessous

$$sim(\omega) = \sum_{i \neq j} \frac{\varepsilon \hat{V}_{\omega}(i) \varepsilon \hat{V}_{\omega}(j)}{\|\varepsilon \hat{V}_{\omega}(i)\| \|\varepsilon \hat{V}_{\omega}(j)\|} \quad (1.53)$$

$$sim(\omega_1) \leq sim(\omega_2) \leq \dots \leq sim(\omega_N) \quad (1.54)$$

2. Pour ω_1 assigner $\hat{V}_{\omega_1}(N_s; i)$ à $\hat{Y}_{\omega_1}(N_s; i)$ comme

$$\hat{Y}_{\omega_1}(N_s; i) = \hat{V}_{\omega_1}(N_s; i), \quad i = 1, \dots, n_s \quad (1.55)$$

3. Pour chaque ω_k trouver une permutation $\sigma_{\omega}(i)$ qui maximise la corrélation entre l'enveloppe de ω_k et l'enveloppe jointe (spectrogrammes divisés) de ω_1 à ω_{k-1} . Cette opération est effectuée en maximisant

$$\sum_{i=1}^n \varepsilon \hat{V}_{\omega_k}(\sigma_{\omega}(i)) \cdot \left(\sum_{j=1}^{k-1} \varepsilon \hat{Y}_{\omega_j}(i) \right) \quad (1.56)$$

pour toute les permutations possibles σ_{ω} avec $i = 1, \dots, n_s$.

4. Assigner la permutation appropriée à $\hat{Y}_{\omega_k}(N_s; i)$

$$\hat{Y}_{\omega_k}(t_s; i) = \hat{V}_{\omega_k}(N_s; \sigma_{\omega}(i)) \quad (1.57)$$

Le résultat est la résolution de la permutation et l'obtention de spectrogrammes séparés définis par

$$\tilde{Y}(\omega, N_s; i) = \tilde{\tilde{Y}}_\omega(N_s; i). \quad (1.58)$$

Les sources reconstruites sont obtenues par la TFFGI des spectrogrammes séparés. Elles ont les mêmes dimensions que les observations et sont définies par

$$\tilde{y}(n; i) = \frac{1}{2\pi} \frac{1}{W(n)} \sum_{N_s} \sum_{\omega} e^{-j\omega(n-N_s)} \tilde{Y}(\omega, N_s; i) \text{ pour } i = 1, \dots, n_s \quad (1.59)$$

où i est l'indice des composantes indépendantes.

1.6 Conclusion

La place occupée par la SS dans le domaine du traitement numérique augmente constamment depuis son introduction dans les années 90. Les statistiques d'ordre 2 ont été pendant plusieurs années les seules à être considérées pour la résolution du problème. Ces statistiques exploitent seulement une partie de la totalité des propriétés statistiques des signaux traités. L'utilisation de statistiques d'ordre supérieur permet d'exploiter en profondeur ces propriétés. Une amélioration de l'exploitation des statistiques implique une amélioration de l'exploitation de la propriété d'indépendance qui se traduit par une augmentation des performances au niveau de la séparation des sources.

Dans la majorité des cas, la SS appartient à la classe « semi-aveugle ». Généralement les caractéristiques des sources et du système de mélange sont connues. Cette connaissance à priori du problème de séparation permet de connaître l'étendu de variation des caractéristiques du système. Ces hypothèses limitent l'étendu du problème et permettent une meilleure prédiction du comportement des algorithmes.

Lorsque les mélanges sont convolutifs, la séparation de source devient une extension de l'égalisation aveugle (identification paramétrique). L'égalisation opère sur un canal tandis que la séparation opère sur plusieurs canaux. Les techniques utilisées sont valables dans l'un ou l'autre des domaines. Le traitement du mélange convolutif est aussi une extension du traitement du mélange instantané dans le sens où les coefficients sont remplacés par des filtres. Les concepts de bases sont les mêmes pour les deux cas.

Lors de la construction du système de séparation, le choix du séparateur et de l'algorithme de mise à jour sont des étapes importantes. Ils déterminent la stabilité du système et son habileté à séparer les sources. La stabilité est très importante pour obtenir un système de SS robuste. Dans la majorité des cas, les réponses devant être inversés sont longues, possédant des durées de plusieurs centaines de millisecondes. L'inversion de ces systèmes à longues réponses impulsionnelles est un sujet en pleine évolution.

Un cas particulier de SS est celui où les systèmes de mélanges sont dynamiques. Les fonctions de transfert entrant dans la création des mélanges sont variables dans le temps. On peut citer comme exemple le cas où deux locuteurs se déplacent par rapport à des microphones fixes. Pour séparer les locuteurs le système doit s'adapter à une vitesse plus rapide que celle de variation du système de mélange. Les mélanges non-linéaires sont aussi un autre cas spécifique.

La SS touche à un ensemble de domaines connexes. Elle rassemble des concepts de toutes parts provenant autant du domaine de la statistique que de l'acoustique. Les applications possibles de la SS sont intéressantes, variées et leur nombre ne cesse de croître.

CHAPITRE 2

LA TRANSFORMÉE EN ONDELETTES

2.1 Introduction

Les transformées permettent de représenter les signaux de différentes façons. La plus connue de toutes, la TF donne le contenu spectral d'un signal donné. Cette transformée est la décomposition du signal analysé sur une base formée de fonctions sinusoïdales. La transformée en ondelettes (TO) permet des décompositions selon d'autres types de fonctions dénommées les ondelettes.

Les coefficients obtenus lors de l'analyse avec la TOC peuvent être interprétés comme le degré de corrélation entre le signal analysé et l'ondelette utilisée [51]. Par comparaison, les coefficients de la TF représentent le niveau de corrélation entre les fonctions sinus cosinus et le signal. L'ensemble sinus cosinus pour la TF et l'ondelette pour la TOC sont les bases ou fonctions de décomposition. Comme pour la TF, la TOC est une projection d'un signal sur un référentiel. La différence réside dans le référentiel, qui dans ce cas, est constitué par les versions dilatées et translatés de l'ondelette d'analyse (voir la section 2.7 pour la théorie sur la TO).

2.2 La transformée en ondelette et la transformée de Fourier à fenêtre glissante

Une conséquence directe de l'utilisation de la TOC est une analyse temporelle fréquentielle différente de celle obtenue par la TFFG. L'échantillonnage du plan temps fréquence est différent. La TOC est une analyse temporelle fréquentielle dont les résolutions varient en fonction de la position sur la plage temps fréquence. Par comparaison, la TFFG est une analyse temporelle fréquentielle dont la résolution est constante pour toute la plage temps fréquence.

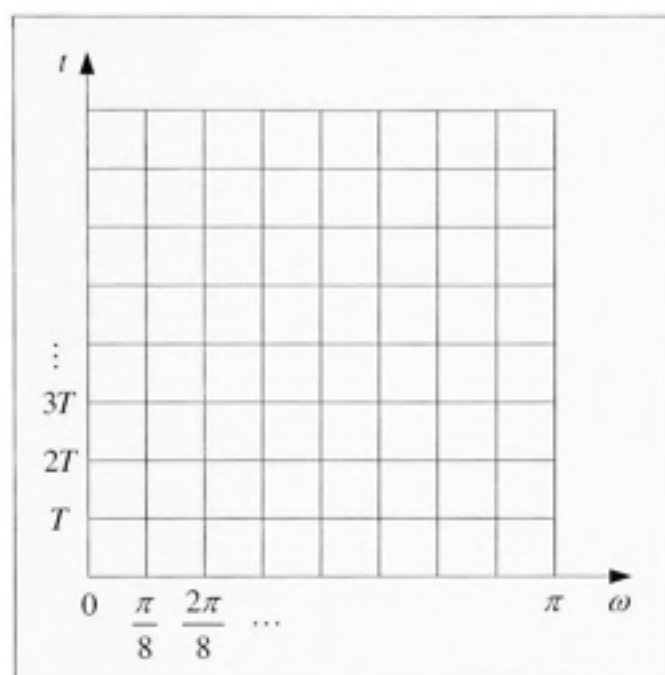


Figure 2.1 *Représentation temps fréquence de la TFFG.*

La TFFG, représentée à la figure 2.1, expose la constance de la résolution avec une uniformité des distances dans l'espace temps fréquence. La TOC, représentée à la figure 2.2, montre la variation de la résolution dans le temps et dans les fréquences. Aux basses fréquences, la résolution temporelle est grande et la résolution fréquentielle petite. Aux hautes fréquences la résolution fréquentielle est grande et la résolution temporelle petite. La résolution fréquentielle est inversement proportionnelle à la résolution temporelle. Un compromis doit donc être fait entre la résolution dans le temps et en fréquence. Cette dualité est connue comme le principe d'incertitude (section 2.6).

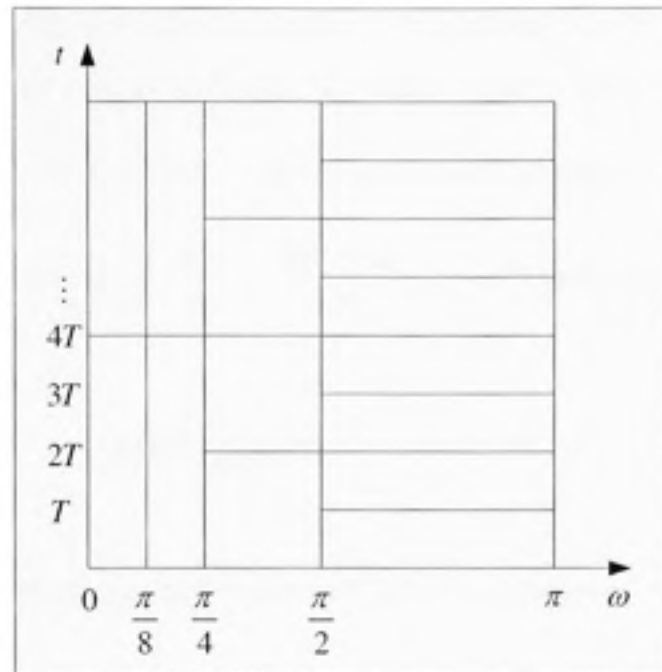


Figure 2.2 *Représentation temps fréquence de la TO.*

2.3 La transformée de Fourier à temps discret et la transformée de Fourier discrète

Avant de définir la TFFG, la transformée de Fourier à temps discret (TFTD) et la TFD doivent être définies. La TFTD est une transformation qui s'applique sur l'ensemble du signal $x(n)$. Le résultat est une représentation de la totalité des composantes fréquentielles contenu dans $x(n)$ sans mettre en évidence leur localisation dans le temps. L'absence d'indice temporel au niveau des coefficients $X(e^{j\omega})$ appuie cet énoncé. La TFTD est définie par

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n}. \quad (2.1)$$

La TFTD $X(e^{j\omega})$ est utilisée pour l'analyse spectrale continue des signaux discrets $x(n)$. La TFD en découle par la discrétisation de $X(e^{j\omega})$ qui est continue. Cette opération est

nécessaire dans le cadre de systèmes discrets. Pour obtenir la TFD, la fréquence ω est échantillonnée sur l'intervalle $[0, 2\pi]$ et la transformée $X(e^{j\omega})$ calculée pour chacune de ces fréquences $\omega_k = \frac{2\pi k}{N}$, $0 \leq k \leq N-1$. Où N représente le nombre de points de la TFD et ω prend les valeurs discrètes $0, \frac{2\pi}{N}, \frac{4\pi}{N}, \dots, \frac{(N-1)2\pi}{N}$. La TFD est définie par

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-j \frac{2\pi}{N} kn}, \quad 0 \leq k \leq N-1. \quad (2.2)$$

La transformation de Fourier discrète inverse (TFDI) est définie par

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{j \frac{2\pi}{N} kn}, \quad 0 \leq k \leq N-1. \quad (2.3)$$

2.4 La transformée de Fourier à fenêtre glissante (TFFG)

La TFFG permet l'analyse temporelle en déterminant la portion du signal analysé avec une fenêtre. Le nouveau paramètre nécessaire est la position temporelle. Ce choix fini de la longueur des données sur lesquelles on applique la TFD nécessite l'introduction d'une fenêtre $v(m)$ dans la transformation. Cette fenêtre permet d'adoucir les transitions aux limites du bloc de données analysé afin de minimiser l'introduction d'artéfacts en fréquence. Les limites du bloc de données sont en fait des transitions rapides qui n'existent pas dans le signal d'origine. Si elles ne sont pas adoucies elles introduiront des composantes fréquentielles inexistantes. La TFFG possède donc deux paramètres, le nombre de points N de la TFD et sa position n . Le fenêtrage du bloc de données sur lequel la TFD est calculée est représenté à la figure 2.3.

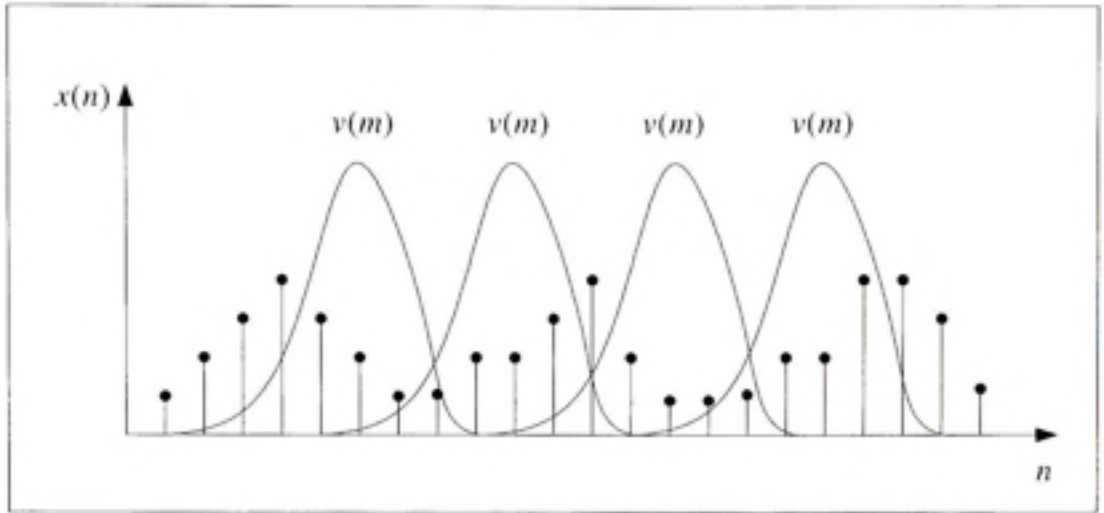


Figure 2.3 Représentation temporelle de la TFFG.

La TFFG est définie par

$$X(k, n) = \sum_{m=0}^{R-1} x(n-m)v(m)e^{-j\frac{2\pi}{N}km}, \quad 0 \leq k \leq N-1 \quad (2.4)$$

où n est l'indice de position, R est la longueur de la fenêtre définie sur $0 \leq m \leq R-1$ et $v(m)$ la fenêtre. La transformée TFFGI est donné par

$$x(n-m) = \frac{1}{Nv(m)} \sum_{k=0}^{N-1} X(k, n)e^{j\frac{2\pi}{N}km}, \quad 0 \leq m \leq R-1. \quad (2.5)$$

Les détails concernant la reconstruction du signal à l'origine de la transformée $X(k, n)$ sont présents dans [52].

2.5 La transformée de Fourier à fenêtre glissante, interprétation en banque de filtres

Cette section a pour objectif d'exposer l'interprétation de la TFFG comme une banque de filtres. Les développements nécessaires à la compréhension en détail de la méthode sont présents dans [53].

La TFFG est équivalente à filtrer le signal avec une banque de filtres passes bandes de largeurs identiques (figure 2.4).

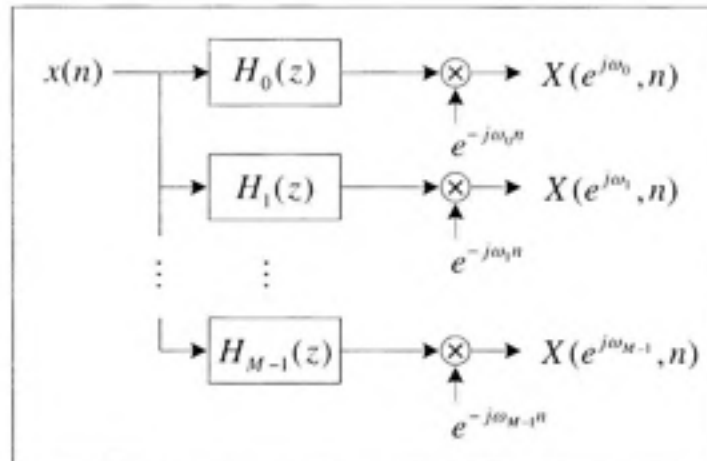


Figure 2.4 *Interprétation en banque de filtres de la TFFG.*

La figure 2.5 montre la réponse des filtres utilisés dans la banque de filtres de la figure 2.4. Les largeurs de bandes sont toutes les mêmes ce qui permet de définir l'analyse avec la TFFG comme une analyse avec une banque de filtres de largeurs uniformes. La résolution est identique peu importe la position en fréquence.

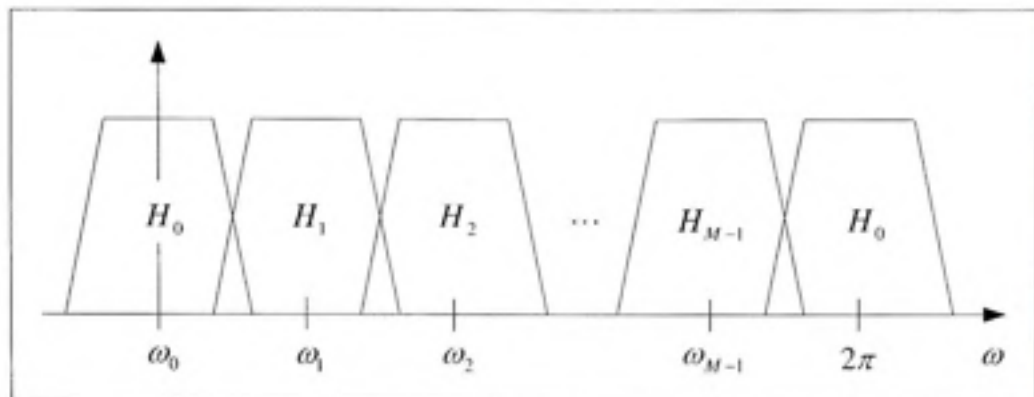


Figure 2.5 *Filtres utilisés dans l'interprétation en banque de filtres de la TFFG.*

2.6 Fenêtre et Incertitude

Les fenêtres sont des fonctions qui permettent de voir une partie précise des données. Elles sont appliquées directement sur les données par le produit scalaire. Les valeurs nulles de la fenêtre éliminent les données, le résultat du produit étant nul à son tour. Les fenêtres les plus communes présentes dans la littérature sont connues sous le nom de : rectangulaire, Gausienne, Hamming, Hanning, Barlett, triangulaire, Blackman et Kaiser. La différence entre les fenêtres réside dans la résolution offerte, chaque fenêtre possède des caractéristiques temporelles fréquentielles différentes. Quelques une des fenêtres citées précédemment sont représentées à la figure 2.6.

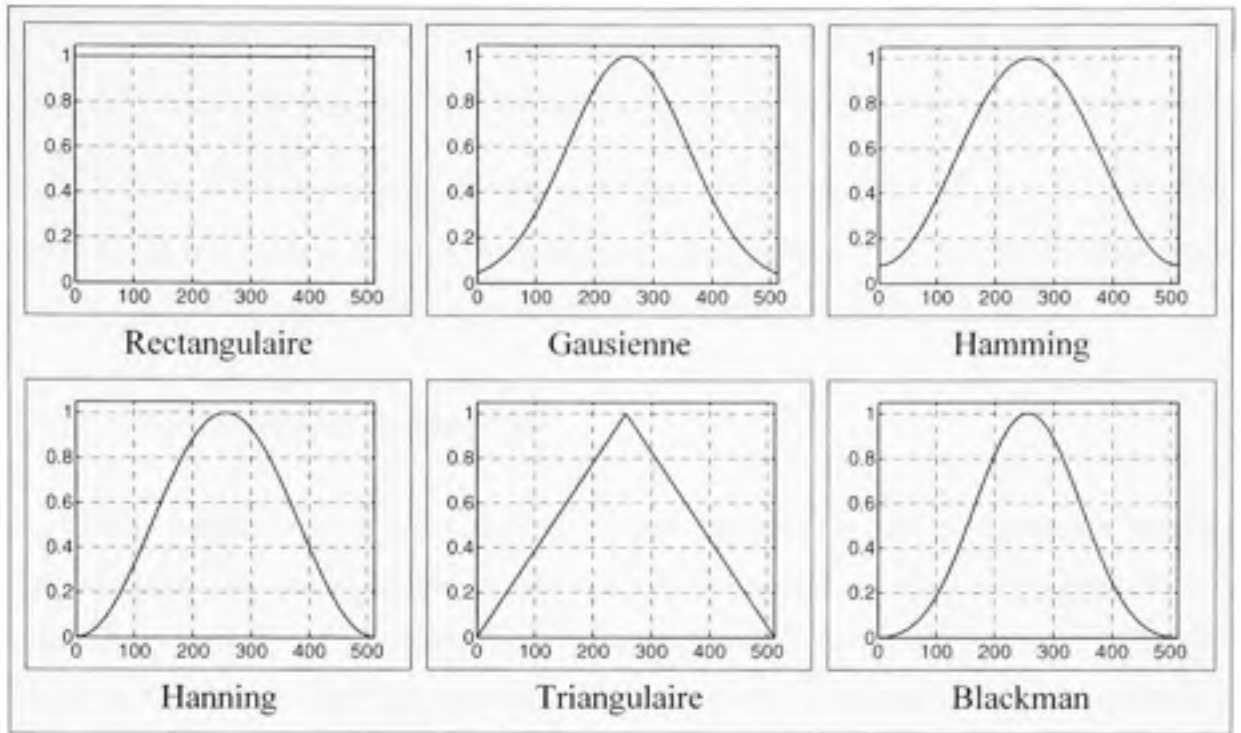


Figure 2.6 Fenêtres couramment utilisées.

La fenêtre optimale est définie par la meilleure résolution possible pouvant être atteinte dans le domaine temporel et fréquentiel. Comme énoncé dans la section 2.2, une fenêtre étroite dans le domaine temporel correspond à une bande passante large en fréquence et vice versa. Un compromis doit être fait entre la résolution dans le temps et en fréquence. Pour déterminer la fenêtre optimale deux quantités sont définies : la durée RMS dans le temps et en fréquence. La durée RMS temporelle est définie par

$$D_t^2 = \frac{1}{E_f} \int_{-\infty}^{\infty} t^2 w_f^2(t) dt \quad (2.6)$$

et la durée RMS fréquentielle par

$$D_f^2 = \frac{1}{2\pi E_f} \int_{-\infty}^{\infty} \Omega^2 |W_f(j\Omega)|^2 d\Omega. \quad (2.7)$$

Les deux quantités sont positives et E est l'énergie de la fenêtre $E_f = \int w_f^2(t) dt$. Le principe d'incertitude stipule que le produit $D_t D_f$ ne peut être arbitrairement petit. Il est limité à $D_t D_f \geq 0.5$ avec égalité si et si seulement $w_f(t) = A_f e^{-\alpha t^2}$ avec $\alpha > 0$. La fenêtre optimale est une fenêtre de forme gaussienne de durée infinie [53]. La TFFG d'un signal avec l'utilisation de cette fenêtre est connue sous le nom de transformée de Gabor.

2.7 La transformée en ondelettes

La TFFG possède une largeur de bande fixe en fonction de la fréquence centrale (interprétation avec une banque de filtres). Cette constance cause une diminution de l'exactitude de la TF en basse fréquence et une augmentation en haute fréquence. Cet effet est dû au nombre de périodes contenues dans la fenêtre d'analyse ou par équivalence à l'uniformité de la banque de filtres [53]. En d'autres termes, une erreur d'évaluation de la TF apparaît lorsque la quantité d'information est insuffisante.

Avec la TO, la largeur du filtre augmente avec la fréquence centrale. Ainsi, aux basses fréquences la fenêtre temporelle est large (étroite en fréquence) et aux hautes fréquences la fenêtre temporelle est étroite (large en fréquence). Lorsque la TO est représentée par une banque de filtres le lien avec l'analyse multirésolution devient évident. L'analyse multirésolution emploie des filtres de largeur et d'amplitudes différentes en fonction de la fréquence. La différence entre la TFFG et la TO réside dans la non-uniformité de l'analyse.

2.7.1 L'ondelette d'analyse

Un grand nombre d'ondelettes d'analyses ont été construites. Plusieurs types ont été rapportés dans la littérature dont certaines portent le nom de leurs créateurs. Les plus connues, représentées à la figure 2.7, sont : Daubechies, Gaussienne, Haar, Mexican Hat, Meyer et Morlet [54].

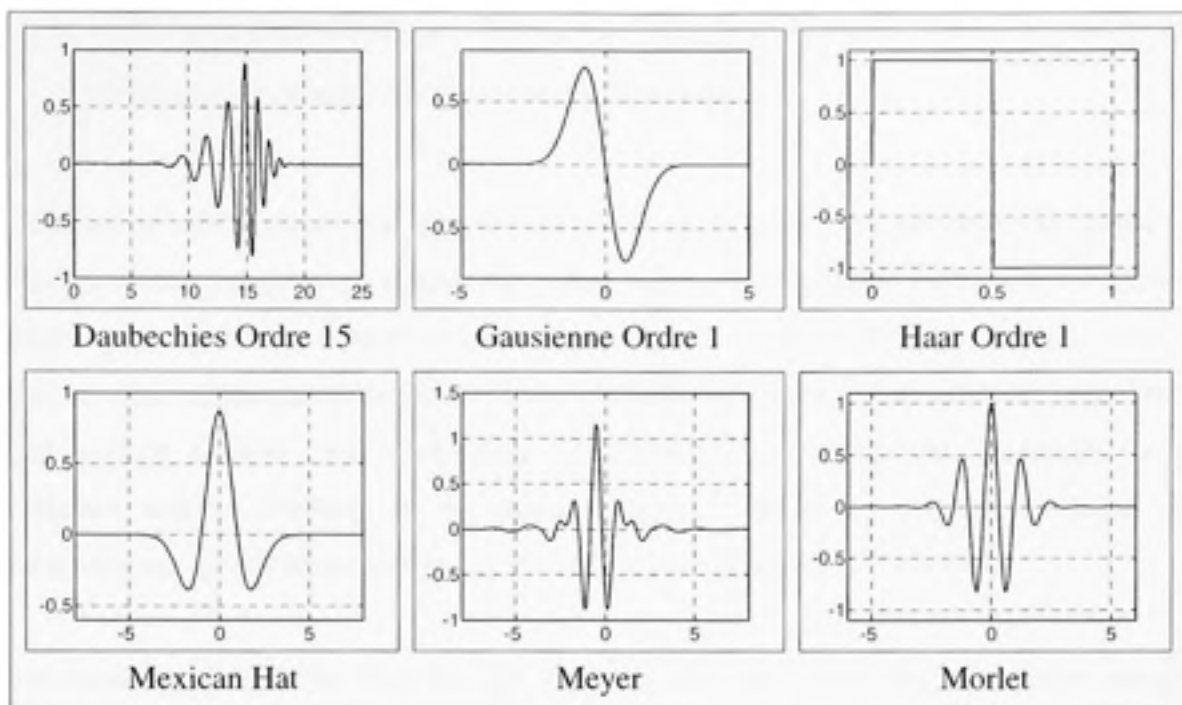


Figure 2.7 *Ondelettes couramment utilisées.*

Le choix de l'ondelette n'est pas complètement arbitraire. Certaines conditions doivent être respectées pour quelle soit admissible [55]. Pour être admissible l'ondelette $\psi(t)$, doit respecter les conditions suivantes :

1. $\psi(t)$ ainsi que son carré doivent être intégrables

$$\int |\psi(t)| dt < \infty \text{ et } \int |\psi(t)|^2 dt < \infty \quad (2.8)$$

2. Si $\hat{\psi}(\omega)$ représente la TF de $\psi(t)$, alors

$$\int \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega < \infty. \quad (2.9)$$

La première condition implique que l'énergie de l'ondelette doit être finie et la seconde que la moyenne de $\psi(t)$ soit nulle ($\hat{\psi}(0) = 0$ où $\int \psi(t)dt = 0$).

Le choix de l'ondelette ne fera pas l'objet d'une analyse dû à sa nature assez complexe. Certaines remarques peuvent quand même être faites. Une multitude d'ondelettes existent chacune possédant des caractéristiques particulières. Les coefficients obtenus lors de l'analyse sont directement reliés au type d'ondelette choisi. Du point de vue de la décomposition linéaire, les coefficients correspondent au degré de similitude ou de corrélation entre l'ondelette et le signal analysé. Ainsi la mise en évidence des caractéristiques du signal analysé par la TO repose sur la forme de l'ondelette.

Il est intéressant de noter que la TO se prête bien à l'étude des transitions rapides. L'ondelette peut être choisie pour détecter les discontinuités car l'ondelette d'analyse oscille. On peut citer comme exemple l'ondelette de Haar. La TF n'effectue pas cette décomposition avec un niveau de performance élevé dû au phénomène de Gibbs [56] qui introduit des dépassements et des oscillations aux discontinuités.

2.7.2 La transformée en ondelettes continues

La TOC utilise les versions dilatées et translatées de l'ondelette mère $\psi(t)$ pour effectuer l'analyse. Avec $\psi(t)$ l'ondelette d'analyse, a le paramètre de dilatation et b le paramètre de translation, $\psi_{a,b}(t)$ est la version dilatée décalée de l'ondelette de base. L'ondelette est transformée dans un plan dilatation translation par

$$\psi_{a,b}(t) = \psi\left(\frac{t-b}{a}\right). \quad (2.10)$$

Pour chaque point (a, b) avec $a > 0$ et $b \in \mathbb{R}$, une ondelette $\psi_{ab}(t)$ est définie. La TOC est le résultat du produit scalaire entre l'ondelette $\psi_{ab}(t)$ et le signal défini par $s(t)$. Le résultat de la transformation est une fonction $z(b, a)$, définie sur la plage dilatation translation par

$$z(b, a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi\left(\frac{t-b}{a}\right) x(t) dt \quad (2.11)$$

Pour chaque point (a, b) une valeur $z(a, b)$ de la TOC de $s(t)$ existe. La TOCI permettant de reconstruire $s(t)$ à partir des coefficients $z(a, b)$ est définie par

$$x(t) = \frac{1}{c_x} \iint \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) z(b, a) \frac{1}{a^2} da db \quad (2.12)$$

où c_x est une constante qui dépend du choix de l'ondelette $\psi(t)$ [55]. Dans le domaine numérique l'intégrale de la formule (2.11) peut être approchée par une convolution et l'intégration finale sur les échelles de la formule (2.12) par une intégration numérique [57]. La méthode trapézoïdale est une des méthodes existantes pouvant être employée pour estimer l'intégration.

2.8 L'analyse multirésolution

L'analyse multirésolution correspond à l'utilisation d'une banque de filtres de largeurs non-uniformes pour l'analyse des signaux. La figure 2.5 montre le module de la réponse en fréquence d'une banque de filtres menant à une décomposition uniforme en sous bandes. La résolution obtenue est la même peu importe la position en fréquence. La figure 2.8 montre une décomposition similaire effectuée cette fois avec une banque de filtres non-uniformes. Cette analyse permet d'obtenir une décomposition de la bande passante du signal en plusieurs bandes de largeur différentes. L'appellation multirésolution provient de cette

variation de la résolution en fonction de la fréquence. L'analyse multirésolution constitue le bloc de construction le plus important pour construire la TOD et la TPO [58].

L'ondelette ψ peut-être construite pour que la famille translatée et dilatée

$$\left\{ \psi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t - 2^j n}{2^j}\right) \right\}_{(j,n) \in \mathbb{Z}^2} \quad (2.13)$$

soit une base orthonormale de $L^2(\mathbb{R})$. Les ondelettes orthogonales à l'échelle 2^j contiennent les variations du signal à la résolution 2^{-j} . La construction de ces bases peut donc être reliée à des approximations multirésolutions d'un signal. Ce lien mène à une équivalence entre les bases d'ondelettes et les filtres miroirs en quadrature.

La définition suivante spécifie les propriétés mathématiques des espaces multirésolutions. Une suite $\{V_j\}_{j \in \mathbb{Z}}$ de sous-espaces fermés $\{V_j\}_{j \in \mathbb{Z}}$ de $L^2(\mathbb{R})$ est une approximation multirésolution si les 6 propriétés suivantes sont vérifiées [59]:

$$\forall (j,k) \in \mathbb{Z}^2, f(t) \in V_j \Leftrightarrow f(t - 2^j k) \in V_j \quad (2.14)$$

$$\forall j \in \mathbb{Z}, V_{j+1} \subset V_j \quad (2.15)$$

$$\forall j \in \mathbb{Z}, f(t) \in V_j \Leftrightarrow f\left(\frac{t}{2}\right) \in V_{j+1}, \quad (2.16)$$

$$\lim_{j \rightarrow +\infty} V_j = \bigcap_{j=-\infty}^{+\infty} V_j = \{0\} \quad (2.17)$$

$$\lim_{j \rightarrow -\infty} V_j = \text{adhérence}\left(\bigcup_{j=-\infty}^{+\infty} V_j\right) = L^2 \quad (2.18)$$

Il existe un θ tel que $\{\theta(t-n)\}_{n \in \mathbb{Z}}$ soit une base de Riesz de V_0 .

où $f \in L^2(\mathbb{R})$, 2^{-j} est une résolution, 2^j est un paramètre d'échelle et V_j est un espace regroupant toutes les approximations possibles pour la résolution 2^{-j} .



Figure 2.8 *Décomposition du spectre par une analyse multirésolution.*

Les figures 2.11 et 2.12 permettent d'obtenir la décomposition en sous bandes de la figure 2.8. Les signaux $u_k(n)$ obtenus à la sortie de la banque de filtres ont les largeurs de bandes exposées à la figure 2.8.

2.8.1 La banque de filtres en quadrature (deux canaux)

La TOD utilise des banques de filtres en quadratures [53] (figure 2.9) pour effectuer l'analyse des signaux. Les équations concernant ce type de structure sont développées dans cette section. Ce type de banque de filtres est à la base de la construction de l'arbre de décomposition et de reconstruction de la TOD et de la TPO.

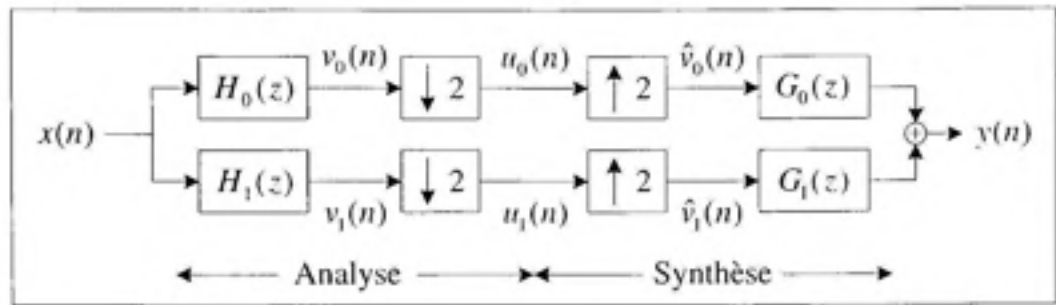


Figure 2.9 Banque de filtres MEQ deux canaux.

La banque de filtres de la figure 2.9 contient les deux filtres d'analyse $H_0(z)$, $H_1(z)$ et les deux filtres de synthèse $G_0(z)$ et $G_1(z)$. Les réponses en fréquences des filtres d'analyse sont montrées à la figure 2.10. Le filtre $H_0(z)$ est un passe bas tandis que le filtre $H_1(z)$ est un filtre passe haut.

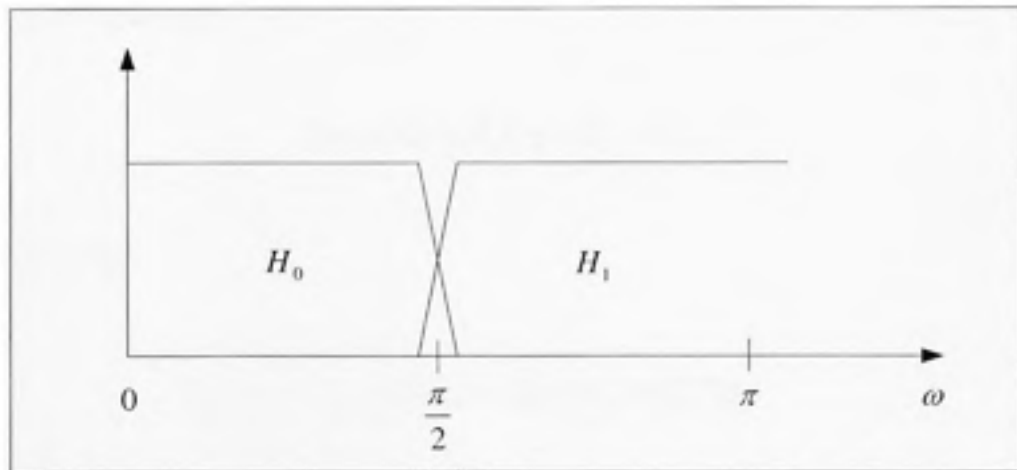


Figure 2.10 Réponses en fréquences banque de filtres en quadrature deux canaux.

La banque de filtres de la figure 2.9 est analysée dans le domaine fréquentiel. Pour cette banque les signaux filtrés sont

$$V_k(z) = H_k(z)X(z), \quad (2.19)$$

après décimation

$$U_k(z) = \frac{1}{2} \left[V_k(z^{1/2}) + V_k(-z^{1/2}) \right]. \quad (2.20)$$

Du côté de la synthèse, les signaux obtenus après interpolation sont

$$\hat{V}_k(z) = U_k(z^2). \quad (2.21)$$

Après manipulation ils deviennent

$$\hat{V}_k(z) = \frac{1}{2} \{V_k(z) + V_k(-z)\} = \frac{1}{2} \{H_k(z)X(z) + H_k(-z)X(-z)\}. \quad (2.22)$$

Le signal reconstruit est

$$Y(z) = G_0(z)\hat{V}_0(z) + G_1(z)\hat{V}_1(z), \quad (2.23)$$

de façon complète

$$Y(z) = \frac{1}{2} \{H_0(z)G_0(z) + H_1(z)G_1(z)\}X(z) + \frac{1}{2} \{H_0(-z)G_0(z) + H_1(-z)G_1(z)\}X(-z). \quad (2.24)$$

Le premier terme de l'équation est connu sous le nom de fonction de transfert de la distorsion et le second comme le terme de recouvrement. Pour obtenir une reconstruction parfaite le terme de recouvrement doit être nul et que la distorsion réduite à un délai avec une amplitude unitaire.

Ces conditions sont respectivement définies par

$$H_0(-z)G_0(z) + H_1(-z)G_1(z) = 0 \quad (2.25)$$

et par

$$H_0(z)G_0(z) + H_1(z)G_1(z) = 2z^{-l}. \quad (2.26)$$

Si les filtres d'analyse et de synthèse sont choisis adéquatement (par exemple les filtres de Daubechies) les banques de filtres MEQ correspondants constituent une analyse du signal en ondelette et sa reconstruction.

2.9 La transformée en ondelettes discrètes

La génération d'un ensemble de sous séquences $u_k(n)$, $0 \leq k \leq L$, à partir d'une séquence d'entrée $x(n)$ utilisant les filtres d'analyse d'une banque de filtres MEQ en octave de L niveaux est connu comme étant la TOD [52]. La figure 2.11 représente une décomposition de 3 niveaux et la figure 2.12 la même décomposition sous la forme d'une banque de filtres 4 canaux. Le filtre $H_{1,1}(z)$ est la transformée en z de l'ondelette mère. Les filtres $H_{2,1}(z)$, $H_{3,1}(z)$, $H_{4,1}(z)$ sont les transformées en z des versions dilatés de l'ondelette mère. Cette représentation met en évidence le lien entre l'analyse multirésolution et la TOD. La banque de filtres MEQ permet de construire l'arbre d'analyse associé à la TOD [60].

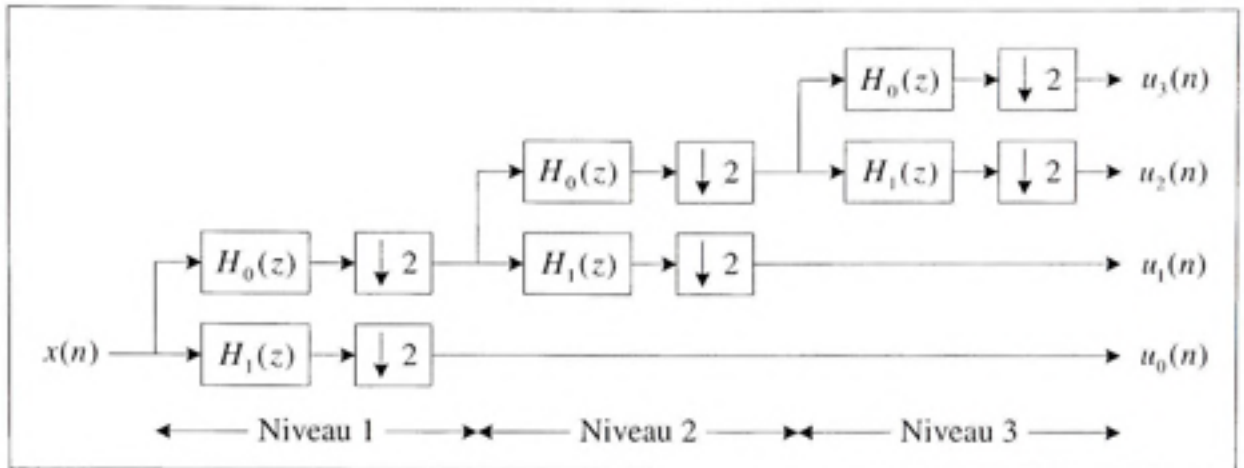


Figure 2.11 Banque de filtres d'analyse en octave de 3 niveaux.

Les fonctions de transfert avant décimation de la banque de la figure 2.11 sont définies par

$$\begin{aligned}
 H_{1,1}(z) &= H_1(z), \\
 H_{2,1}(z) &= H_0(z)H_1(z^2), \\
 H_{3,1}(z) &= H_0(z)H_0(z^2)H_1(z^4), \\
 H_{3,0}(z) &= H_0(z)H_0(z^2)H_0(z^4).
 \end{aligned} \tag{2.27}$$

De façon générale les fonctions de transfert sont définies par

$$H_{k,1}(z) = \left[\prod_{i=0}^{k-2} H_0(z^{2^i}) \right] H_1(z^{2^{k-1}}), \tag{2.28}$$

$$H_{k,0}(z) = \prod_{i=0}^{k-1} H_0(z^{2^i}). \tag{2.29}$$

Les signaux de sorties d'une banque de filtres L niveau en octave MEQ avant décimation sont définis dans le domaine z par

$$X_k(z) = H_{k,1}(z)X(z), \quad 0 \leq k \leq L-1, \quad (2.30)$$

$$X_k(z) = H_{L,0}(z)X(z), \quad k = L, \quad (2.31)$$

et dans le domaine temporel par

$$x_k(n) = \sum_{m=-\infty}^{\infty} h_{k,1}(n-m)x(m), \quad 0 \leq k \leq L-1, \quad (2.32)$$

$$x_k(n) = \sum_{m=-\infty}^{\infty} h_{k,0}(n-m)x(m), \quad k = L. \quad (2.33)$$

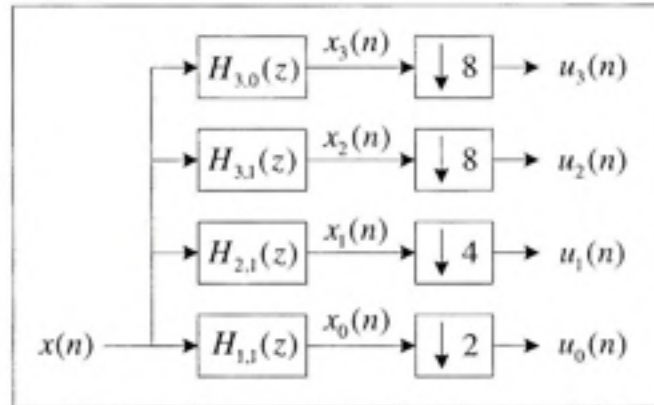


Figure 2.12 *Système équivalent 4 canaux de la figure 2.11.*

La TOD est obtenue après la décimation des signaux à la sortie des filtres. Elle est définie par

$$u_k(n) = \sum_{m=-\infty}^{\infty} h_{k,1}(2^{k+1}n-m)x(m), \quad 0 \leq k \leq L-1, \quad (2.34)$$

$$u_k(n) = \sum_{m=-\infty}^{\infty} h_{k,0}(2^k n-m)x(m), \quad k = L. \quad (2.35)$$

De façon inverse, la reconstruction de la séquence $y(n)$, qui est une réplique de $x(n)$, à partir des séquences $u_k(n)$ pour $0 \leq k \leq L$ en utilisant une banque de filtres de synthèses MEQ en octave de L niveaux pour une reconstruction parfaite est connue comme la transformée en ondelettes discrètes inverse (TODI) [52].

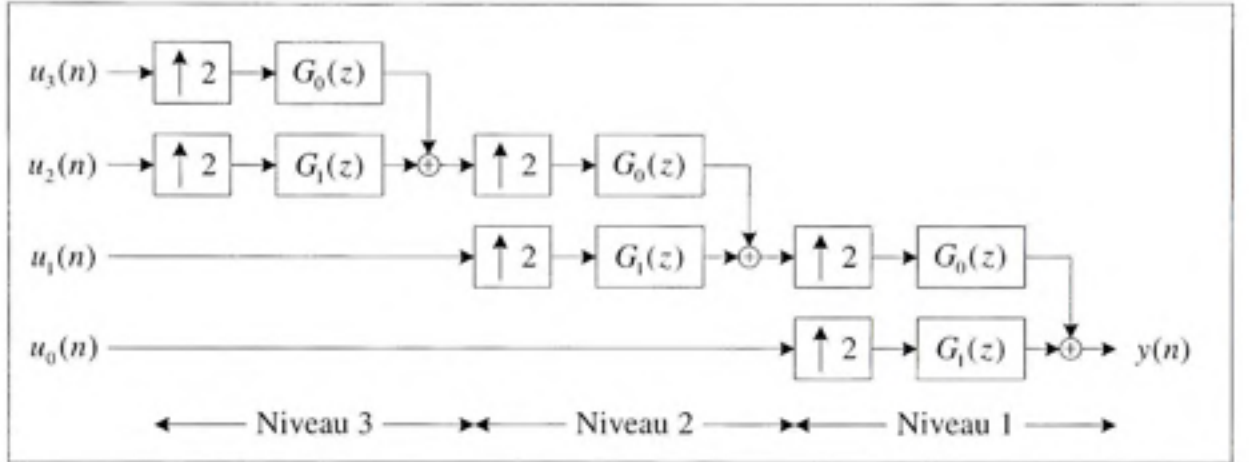


Figure 2.13 Banque de filtres de synthèse en octave de 3 niveaux.

Pour la reconstruction de la figure 2.13 les fonctions de transfert après interpolation sont données par

$$G_{1,1}(z) = G_1(z),$$

$$G_{2,1}(z) = G_0(z)G_1(z^2),$$

$$G_{3,1}(z) = G_0(z)G_0(z^2)G_1(z^4),$$

$$G_{3,0}(z) = G_0(z)G_0(z^2)G_0(z^4). \quad (2.36)$$

Ces fonctions de transfert sont généralisées par

$$G_{k,1}(z) = \left[\prod_{i=0}^{k-2} G_0(z^{2^i}) \right] G_1(z^{2^{k-1}}), \quad (2.37)$$

$$G_{k,0}(z) = \prod_{i=0}^{k-1} G_0(z^{2^i}). \quad (2.38)$$

Si les conditions pour une reconstruction parfaite sont respectées $y(n) = x(n)$. Le signal de sortie $y(n)$ exprimé en termes des signaux de sortie $u(n)$ de la banque de filtres d'analyse est défini par

$$Y(z) = G_{1,1}(z)U_0(z^2) + G_{2,1}(z)U_1(z^2) + \dots + G_{L,1}(z)U_{L-1}(z^{2^L}) + G_{L,0}(z)U_L(z^{2^L}). \quad (2.39)$$

Dans le domaine temporel les sorties sont exprimées par

$$y(n) = \sum_{k=0}^{L-1} \sum_{m=-\infty}^{\infty} g_{k,1}(n - 2^{k+1}m)u_k(m) + \sum_{m=-\infty}^{\infty} g_{L,0}(n - 2^L m)u_L(m) \quad (2.40)$$

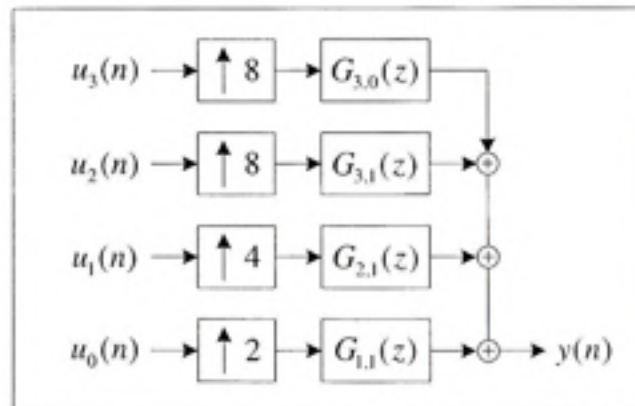


Figure 2.14 Système équivalent 4 canaux de la figure 2.13.

La TOD est construite par l'insertion d'une banque de filtres MEQ de deux canaux dans la branche supérieure d'une autre banque de filtres identique. L'opération est répétée pour l'ajout de chaque niveau supplémentaire. La branche supérieure fournit les approximations et celle inférieure les détails. À la figure 2.11 les détails sont $u_0(n)$ et les approximations, $u_1(n)$, $u_2(n)$ et $u_3(n)$. Les détails sont véhiculés par les hautes fréquences et les approximations par les basses. La décomposition d'une image permet de bien représenter cette appellation. Les contours (détails) proviennent d'une variation rapide du signal (hautes fréquences). Les régions uniformes (approximations) proviennent de variations lentes (basses fréquences) du signal. Le filtre $H_0(z)$ est un passe bas qui fournit les approximations et $H_1(z)$ un passe haut qui fournit les détails.

Un avantage de cette structure est l'utilisation des mêmes filtres d'analyses et de synthèses peu importe le nombre de niveau de décomposition. Une fois les filtres $H_0(z)$, $H_1(z)$, $G_0(z)$ et $G_1(z)$ conçus, l'ensemble de l'arbre est facilement réalisable. Cette structure permet l'obtention d'une analyse en banque de filtres de largeurs inégales ou comme cité précédemment multirésolution. Ce sont les décimations qui permettent de modifier la largeur de bande de chaque branche, les filtres étant toujours les mêmes avec une largeur de bande fixe.

La reconstruction de $x(n)$ est parfaite si la banque de filtres MEQ deux bandes (bloc de base) permet une reconstruction parfaite [52]. La conception de la structure complète est donc équivalente à la conception d'une banque de filtres MEQ de deux canaux.

2.10 La transformée en paquets d'ondelettes

Comme la TOD, la TPO utilise une banque de filtres MEQ comme bloc élémentaire [59]. La différence provient de l'introduction, dans chaque branche de la banque MEQ 2 canaux, d'une banque de filtres MEQ identique. En d'autres termes, pour chaque niveau ajouté, les détails et les approximations sont à leurs tours décomposés en détails et en approximations.

Cette approche donne l'arbre dyadique de la figure 2.15. Cette construction est présente dans la littérature sous le nom de structure MEQ décimée de façon maximale. La figure 2.17 est la représentation classique de la décomposition détaillée de la figure 2.15. Cette représentation est plus compacte et définit plus clairement l'analyse.

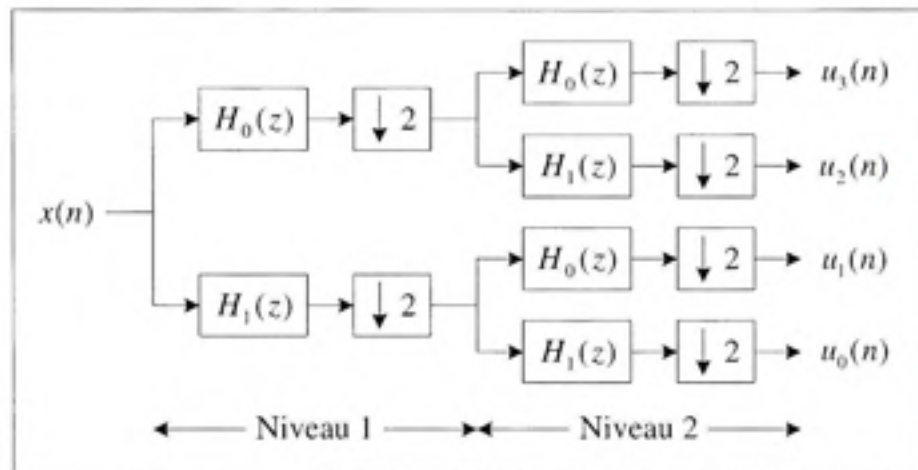


Figure 2.15 Analyse en paquet d'ondelettes.

Les filtres d'analyses équivalents pour la représentation de la figure 2.16 sont définis par

$$\begin{aligned}
 H_{00}(z) &= H_0(z)H_0(z^2), \\
 H_{01}(z) &= H_0(z)H_1(z^2), \\
 H_{10}(z) &= H_1(z)H_0(z^2), \\
 H_{11}(z) &= H_1(z)H_1(z^2).
 \end{aligned} \tag{2.41}$$

La TPO est équivalente à une analyse avec une banque de filtres de largeurs uniformes.

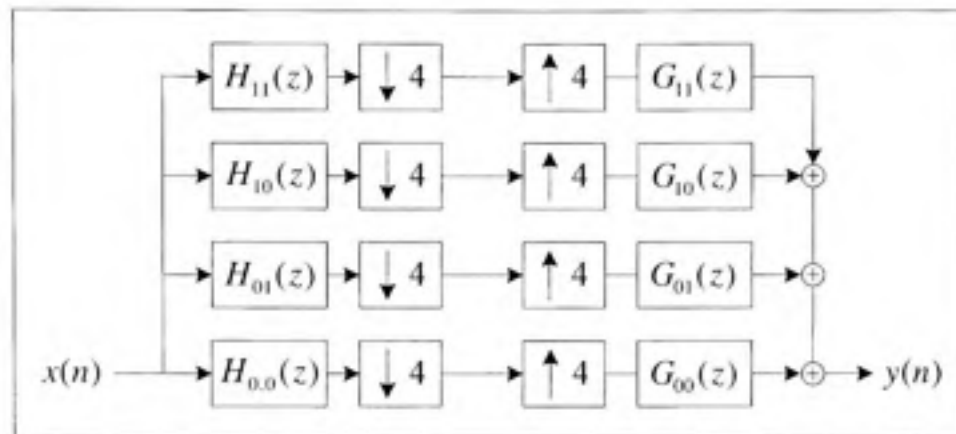


Figure 2.16 Représentation équivalente de l'analyse et de la synthèse avec la TPO pour les figures 2.15 et 2.16.

La transformée en paquets d'ondelettes inverse (TPOI) est obtenue, comme dans le cas de la TOD, par l'inversion du système d'analyse. Le résultat est l'obtention de la structure inverse de la figure 2.17. Elle met en évidence les filtres de synthèse $G_0(z)$ et $G_1(z)$. La conception est du même niveau que la TOD dans le sens où une seule banque de filtres MEQ est à la base de la TPO indépendamment du niveau de décomposition. La différence réside dans le type d'analyse obtenue. Chaque nœud est décomposé ce qui a pour effet d'augmenter la résolution. Avec la TOD les détails obtenus occupent une largeur de bande plus large qu'avec la TPO. Avec la TPO les détails sont décomposés à leur tour. Pour le même niveau de décomposition, le nombre de bandes ou de bases obtenues avec la TPO est supérieur à la TOD.

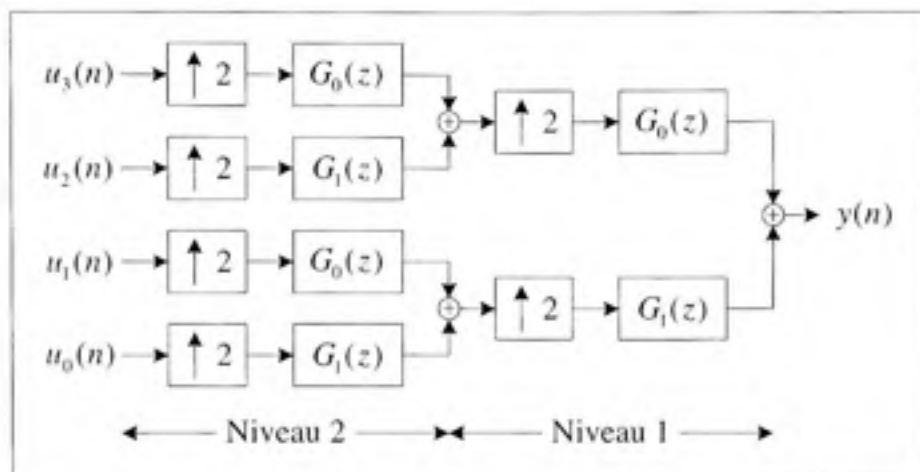


Figure 2.17 *Synthèse en paquet d'ondelettes.*

Les filtres de synthèses équivalents pour la représentation de la figure 2.16 sont définis par

$$\begin{aligned}
 G_{00}(z) &= G_0(z)G_0(z^2), \\
 G_{01}(z) &= G_0(z)G_1(z^2), \\
 G_{10}(z) &= G_1(z)G_0(z^2), \\
 G_{11}(z) &= G_1(z)G_1(z^2).
 \end{aligned} \tag{2.42}$$

2.10.1 Choix du meilleur arbre de décomposition

La TPO possède l'avantage de donner plusieurs bases à partir desquelles il est possible d'effectuer la reconstruction [61]. L'approche généralement employée pour choisir la meilleure décomposition se base sur l'utilisation de l'entropie comme critère de sélection. Par exemple, pour chaque nœud de l'arbre de la figure 2.18, l'entropie [59] est calculée. Dépendamment du résultat obtenu, la décomposition s'arrête ou se poursuit. L'arbre final obtenu correspond à une décomposition optimale dictée par le critère utilisé. Le choix de l'arbre de décomposition permet, en outre, d'obtenir uniquement les données qui sont utiles à une analyse spécifique.

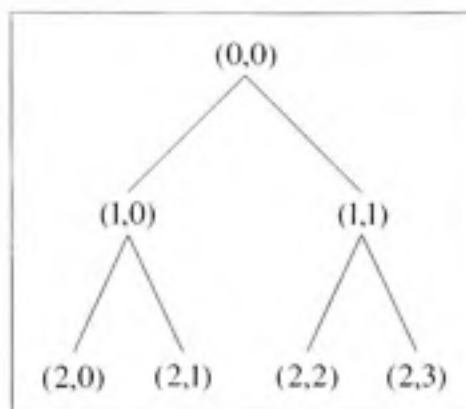


Figure 2.18 *Arbre de décomposition
2 niveaux.*

Les critères les plus couramment utilisés sont : Shannon, Sure, énergie logarithmique et énergie normalisée. Bien entendu, il est possible de développer des critères personnels en fonction du type d'analyse effectuée.

2.11 Applications de la transformée en ondelettes

Le nombre d'applications possibles de la TO ne cesse d'augmenter. Actuellement, elle fait partie intégrante du domaine du traitement des signaux. Le nombre de documents sur le sujet ne cesse d'augmenter. L'ensemble du domaine ne peut être parcouru pour des raisons évidentes. Les domaines d'applications les plus connus sont [56, 62, 63] : le son, l'acoustique, l'imagerie, la géophysique, les fractals, la turbulence, la spectroscopie et la médecine.

Parmi les applications possibles, les plus connues sont : le débruitage [58], la compression [64], l'extraction de caractéristiques [58] et l'étude des transitions [58]. Cette liste permet de constater la place occupée par la TO.

Au-delà des grandes lignes citées précédemment, on peut citer les exemples concrets suivants : la compression d'image avec différents formats [65] dont JPEG2000, la

compression vidéo [66] avec des codecs comme SNOW et BINK, la reconnaissance de formes [67].

2.12 Conclusion

L'analyse en ondelette permet une décomposition selon différentes bases autre que celle utilisées par l'analyse de Fourier. Elle convient particulièrement à l'étude des signaux non stationnaires. Sa grande flexibilité de résolution temporelle et fréquentielle lui permet de représenter plus précisément les différentes bandes de fréquence de ces signaux. La TOD et la TPO sont basées sur l'utilisation d'un bloc élémentaire, la banque de filtres MEQ 2 canaux. De son côté la TOC est équivalente au filtrage du signal par l'ondelette.

L'utilisation d'une ondelette spécifique permet l'extraction de caractéristiques particulières du signal. Un bon exemple est l'extraction de caractéristiques pour la reconnaissance de formes. Comme énoncé précédemment, les coefficients de la TO peuvent être vus comme une corrélation entre l'ondelette et le signal. Les caractéristiques recherchées peuvent donc être utilisées pour construire l'ondelette et vice versa.

La TPO est un outil très intéressant dû au nombre de bases offertes pour la reconstruction du signal analysé. Cette variété fait apparaître les critères de décomposition pour l'obtention de l'arbre optimal. Comme pour l'ondelette d'analyse, les critères de décompositions peuvent être développés par l'analyste ou sélectionnés dans la liste présente dans la littérature. Une vaste gamme de possibilités est offerte dû à la flexibilité du choix de l'ondelette et de l'arbre.

CHAPITRE 3

LA SÉPARATION DE SOURCES À L'AIDE DE LA TRANSFORMÉE EN ONDELETTES CONTINUES

3.1 Introduction

Dans ce chapitre, les éléments essentiels de la méthode développée dans le cadre de ce travail sont exposés. Cette nouvelle méthode utilise la TOC [68] comme première étape dans l'algorithme de SS. Cette transformée ne semble pas avoir été considérée dans le développement initial d'un algorithme de SS.

Quelques méthodes utilisant la TOD et la TPO [10, 42-47] comme prétraitement sont présentées dans la littérature. Avec la TOD et la TPO, les décimations de l'arbre d'analyse diminuent le nombre de délais devant être compensés par l'algorithme sans toutefois les éliminer complètement. Les conclusions présentées dans la littérature montrent une amélioration des performances due à la diminution des longueurs des fonctions de transfert devant être inversés [10].

La TOD et la TPO génèrent des coefficients ayant une distribution différente de celle des signaux sources. Dépendamment des distributions recherchées, l'impact sur les performances peut être positif. Une distribution se rapprochant de celle correspondant à une solution accélère la convergence de l'algorithme de séparation [10].

L'utilisation de banque de filtres de largeurs uniformes est reconnue comme augmentant les performances des systèmes de séparation [49]. La TPO et la TOD sont des décompositions avec banques de filtres non-uniformes. Les performances de SS sont donc affectées par l'utilisation de ce type d'analyse.

L'algorithme de séparation au cœur du système de séparation diffère d'une méthode à l'autre. Dans la majorité des cas, la méthode d'optimisation utilisée est celle du gradient naturel. D'autres méthodes comme celles exposées au chapitre 1 peuvent être utilisées dans le bloc de séparation. Dans la méthode proposée, l'algorithme de séparation EASI est choisi (section 1.5.3).

La méthode de SS proposée est développée pour un mélange de type convolutif. Comme vu au chapitre 1, le traitement du mélange convolutif nécessite l'utilisation d'une matrice de filtres dans la réalisation du séparateur. La TOC n'élimine pas les effets du mélange convolutif.

La séparation des coefficients obtenus avec la TOC est donc effectuée dans le domaine fréquentiel où une matrice de séparation W_N de coefficients permet de séparer chaque bande de fréquence obtenue avec la TFFG des coefficients d'ondelettes. Cette méthode de séparation suppose une matrice de mélange A constituée de filtres à phases minimales. L'algorithme de séparation ne possède pas une structure permettant la séparation d'un mélange provenant d'une matrice de mélange A constituée de filtres à phases non-minimales (voir section 1.3.2).

3.2 Structure du séparateur

La première étape est l'analyse de chaque mélange avec la TOC. Les coefficients obtenus sont séparés par un bloc de séparation en fréquence utilisant la TFFG et l'algorithme de séparation EASI. La synthèse des coefficients séparés obtenus permet d'obtenir les sources reconstruites.

Comme cité dans la section 3.1 le séparateur doit être en mesure d'éliminer les effets introduits par la matrice de mélange A . Le bloc en fréquence permet de traiter ces effets en remplaçant les convolutions par des multiplications. Cette propriété permet de séparer chaque bande avec l'utilisation d'une matrice de coefficients.

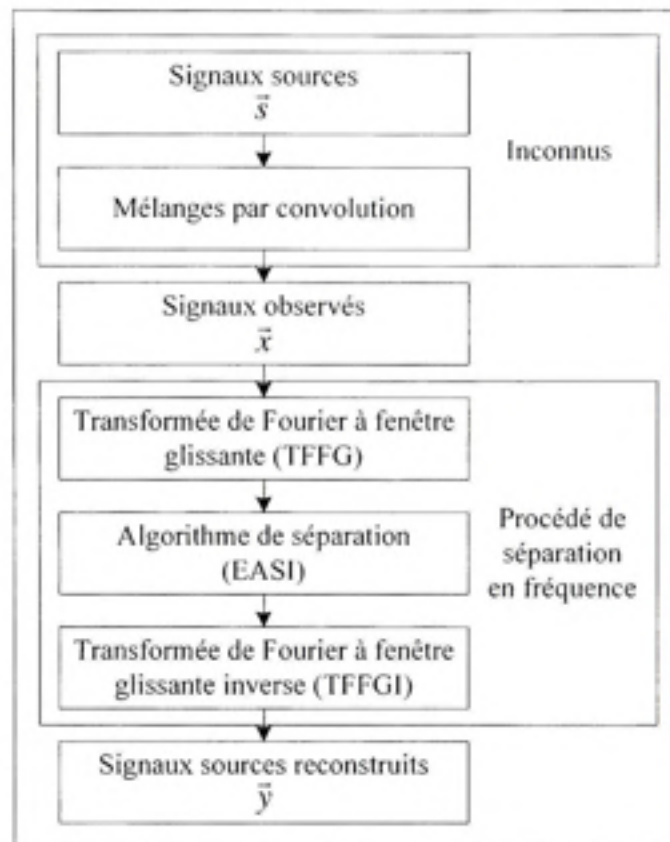


Figure 3.1 *Algorithme de SS en fréquence.*

La figure 3.1 est la représentation en bloc d'un algorithme de séparation en fréquence constitué par la TFFG et l'algorithme EASI. Ce dernier constitue le cœur du procédé de séparation. L'algorithme proposé applique ce bloc de séparation aux coefficients de la TOC utilisée comme prétraitement des mélanges.

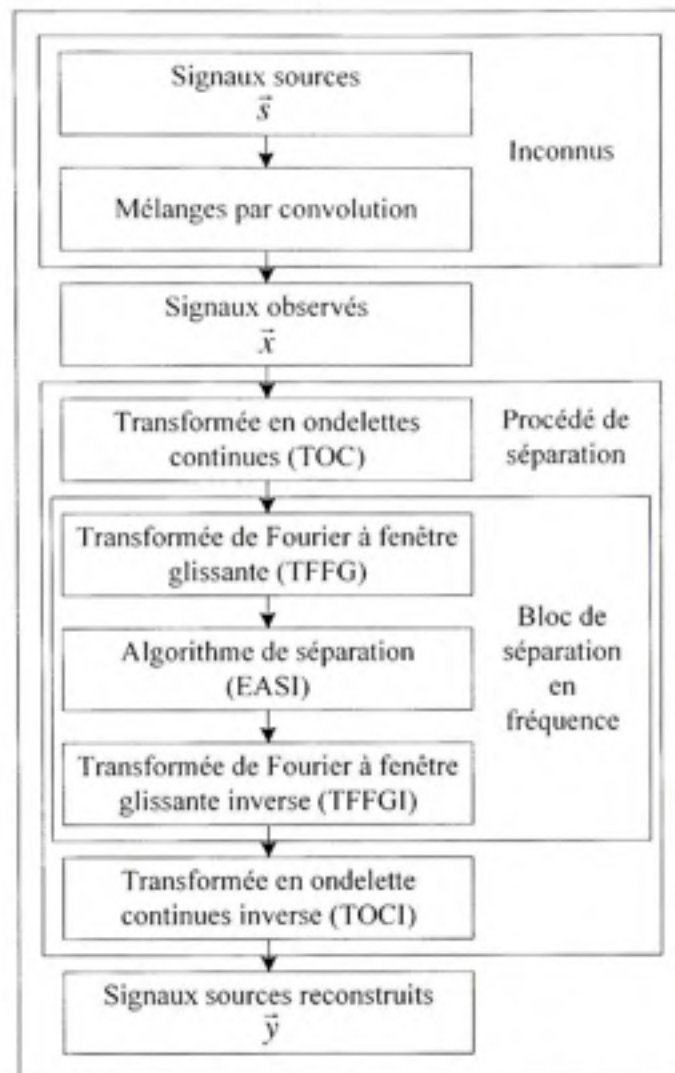


Figure 3.2 *Algorithme de SS avec la TOC et la TFFG.*

3.3 La séparation de sources avec la transformée en ondelettes continues

La figure 3.3 représente en détail les étapes de la SS en fréquence avec la résolution des problèmes de permutation et de dilatation. Le problème de dilatation provient de l'impossibilité pour l'algorithme de déterminer l'amplitude des sources originales. Le problème de permutation provient du mélange de l'appartenance des bandes à chaque source en fréquence.

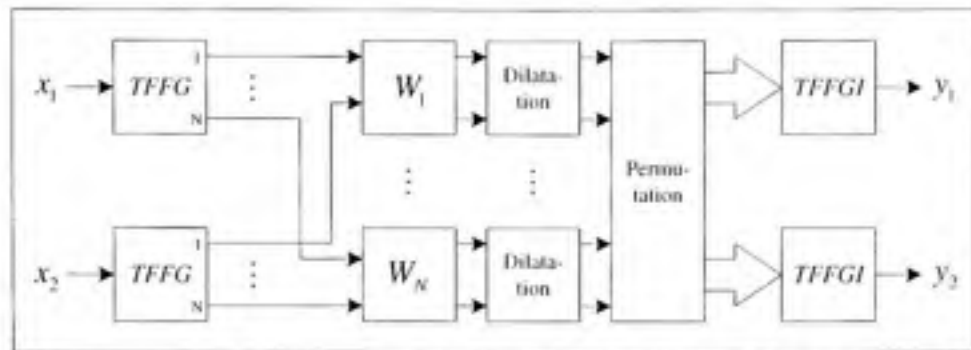


Figure 3.3 Représentation détaillée de la SS en fréquence.

La figure 3.4 représente en détail le séparateur proposé. Le bloc de séparation de la figure 3.3 est utilisé pour séparer les coefficients obtenus pour chaque échelle de la TOC. Le problème de la dilatation n'a pas à être résolu pour chaque échelle car il a été résolu en fréquence. Le problème de permutation par contre doit être résolu au niveau des composantes fréquentielles et des coefficients d'ondelettes. L'élimination des permutations en fréquence n'implique pas l'élimination des permutations aux niveaux des échelles. Une échelle reconstruite après sa séparation dans le domaine fréquentiel peut appartenir à une des sources devant être reconstruite. L'appartenance de chaque échelle à sa source d'origine doit être déterminée afin de reconstruire correctement les sources dans le domaine temporel. Ce problème de permutation au niveau des échelles doit être résolu avant la synthèse finale avec la TOCI.

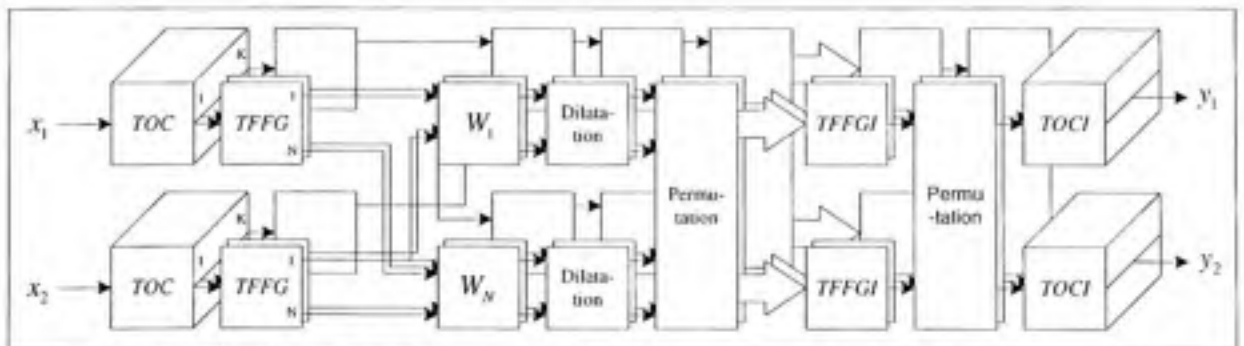


Figure 3.4 Représentation détaillée de la SS avec la TOC.

La méthode de la figure 3.4 est définie étape par étape comme suit :

1. Analyse de chaque mélange avec la TOC.
2. Pour chacune des échelles
 - a. Analyse des coefficients avec la TFFG.
 - b. Pour chaque bande spectrale obtenir une matrice de séparation avec l'algorithme EASI
 - c. Résoudre les problèmes de dilatation et de la permutation en fréquence
 - d. Séparation des coefficients pour chaque échelle
 - e. Application de la TFFGI
3. Résoudre la permutation au niveau des échelles
4. Effectuer la synthèse des coefficients pour obtenir les sources reconstruites avec la TOCI

3.3.1 Analyse avec la transformée en ondelettes continues

L'analyse des mélanges avec la TOC définie par

$$\bar{z}(a,b) = \int_{-\infty}^{\infty} \bar{x}(t) \psi_{a,b}^*(t) dt \quad (3.1)$$

avec la dilatation et la translation

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right). \quad (3.2)$$

L'ondelette de Morlet est définie par

$$\psi(t) = e^{j\omega_0 t} e^{-\frac{\omega_0^2 t^2}{2}} \quad (3.3)$$

où ω_0 est la fréquence de l'ondelette.

3.3.2 Choix des échelles

La difficulté avec l'utilisation de la TOC réside dans la synthèse. Un nombre fini d'échelles devrait être choisi; idéalement le plus efficacement possible. Les échelles sont choisies pour obtenir une bande passante correspondant approximativement à la bande téléphonique. Cette largeur de bande correspond à la partie exprimée ou du spectre de la parole. La bande choisie contient les fréquences 300 à 3400 Hertz.

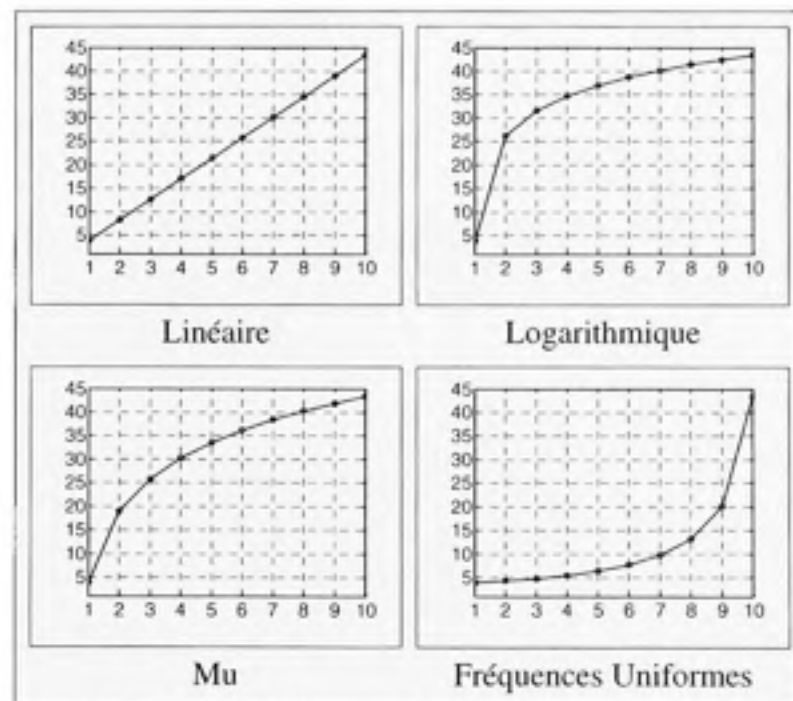


Figure 3.5 Loïs de sélection des échelles.

La bande choisie, le nombre et la répartition des échelles doivent être définies. Chaque loi contrôle la répartition et le nombre d'échelles. La figure 3.5 représente chacune de ces lois pour un nombre de 10 échelles. L'abscisse représente l'indice de l'échelle entre 1 et 10 tandis que l'ordonnée représente l'échelle. L'échelle maximale a_{\max} et minimale a_{\min} correspondent respectivement aux pseudos-fréquences 300 et 3400 Hz.

Les quatre lois de sélection : linéaire, logarithmique, mu et fréquences uniformes, sont respectivement définies par

$$a_{\text{lin}}(k) = a_{\min} + (k-1) \left(\frac{a_{\max} - a_{\min}}{K-1} \right), \quad (3.3)$$

$$a_{\log}(k) = \mu_b \log \left[e^{\frac{a_{\min}}{\mu_b}} + (k-1) \left(\frac{e^{\frac{a_{\max}}{\mu_b}} - e^{\frac{a_{\min}}{\mu_b}}}{K-1} \right) \right], \quad (3.4)$$

$$a_{\mu}(k) = a_{\min} + a_{\max} \frac{x_{\mu}}{|x_{\mu}|} \left(\frac{\log \left(\frac{1 + \mu_o |x|}{a_{\max}} \right)}{\log(1 + \mu_o)} \right), \quad (3.5)$$

$$a_{\text{mi}}(k) = \frac{f_c}{k \left(\frac{f_{\rho \max} - f_{\rho \min}}{N-1} \right)}, \quad (3.6)$$

où

$$x_{\mu} = (k-1) \left(\frac{a_{\max} - a_{\min}}{N-1} \right). \quad (3.7)$$

Pour toutes les lois, l'index de l'échelle k est donné par

$$k = 1, 2, \dots, N. \quad (3.8)$$

Les échelles minimale a_{\min} et maximale a_{\max} sont respectivement données par

$$a_{\min} = \frac{f_c}{f_{p\max}}, \quad a_{\max} = \frac{f_c}{f_{p\min}}. \quad (3.9)$$

où $f_{p\min}$ est la pseudo-fréquence minimale, $f_{p\max}$ est la pseudo-fréquence maximale et f_c la fréquence centrale de l'ondelette. Pour la loi mu et logarithmique μ_a et μ_b sont des paramètres de forme. La pseudo-fréquence minimale correspond à la fréquence minimale et la pseudo-fréquence maximale à la fréquence maximale choisie pour couvrir la largeur de bande désirée. Par exemple pour le signal de la parole $f_{p\min} = 300\text{ Hz}$ et $f_{p\max} = 3400\text{ Hz}$.

3.3.3 Bloc de séparation en fréquence

Le bloc de séparation en fréquence permet d'effectuer la séparation des coefficients de la TOC pour chaque échelle. Il est constitué des éléments représentés à la figure 3.1 soit : la TFFG, l'algorithme EASI et la TFFGI.

L'analyse avec la TFFG est définie par

$$\hat{\tilde{Z}}(\omega, a, N_s) = \sum_n e^{-j\omega n} \tilde{z}(a, n) w_f(n - N_s) \quad (3.10)$$

avec

$$\omega = 0, \frac{1}{N}2\pi, \dots, \frac{N-1}{N}2\pi, \quad (3.11)$$

et

$$N_s = 0, \Delta N, 2\Delta N, \dots \quad (3.12)$$

où ω est la fréquence, N le nombre de point de la TFD, N_s la position de la fenêtre, $w_f(\cdot)$ la fenêtre utilisée et ΔN_s le déplacement de la fenêtre. Dans la formule (3.10) l'indice de temps discret n prend la place de la translation b dans le vecteur des coefficients $\bar{z}(a, b)$ pour donner le nouveau vecteur $\bar{z}(a, n)$.

La fenêtre de Hamming est définie par

$$w_f(n+1) = 0.54 - 0.46 \cos\left(2\pi \frac{n_f}{l_f - 1}\right) \quad (3.13)$$

avec

$$n_f = 0, \dots, l_f - 1. \quad (3.14)$$

La séparation des bandes spectrales avec l'algorithme EASI normalisé

$$W_{i+1} = W_i - u \left[\frac{\hat{\bar{Z}}\hat{\bar{Z}}^T - I}{1 + \mu\hat{\bar{Z}}^T\hat{\bar{Z}}} + \frac{g(\hat{\bar{Z}})\hat{\bar{Z}}^T - \hat{\bar{Z}}g(\hat{\bar{Z}})^T}{1 + \mu|\hat{\bar{Z}}^T g(\hat{\bar{Z}})|} \right] W_i \quad (3.15)$$

avec

$$W_i = W_i(\omega, a), \quad \hat{\bar{Z}} = \hat{\bar{Z}}(\omega, a, N_s) \quad (3.16)$$

où W_{i+1} est la matrice de séparation mise à jour, $g(\cdot)$ une non-linéarité choisie, I la matrice identité et μ le pas itératif.

Après la séparation de chacune des composantes spectrales, les problèmes de dilatation de permutation doivent être résolus.

Les composantes spectrales séparées sont

$$\hat{\bar{U}}_{\omega}(a, N_s) = W_i(\omega, a) \hat{\bar{Z}}(\omega, a, N_s) \quad (3.17)$$

et le problème de dilatation est résolu par l'application de l'inverse de la matrice de séparation aux composantes spectrales $\hat{\bar{U}}_{\omega}(a, N_s)$ comme ci-dessous

$$\hat{\bar{V}}_{\omega}(a, N_s; i) = W_i(\omega, a)^{-1} \begin{pmatrix} 0 \\ \vdots \\ \hat{U}_{i,\omega}(a, N_s) \\ \vdots \\ 0 \end{pmatrix}, \quad (3.18)$$

où i signifie la dépendance des TFFG sur ω pour les composantes i de $\hat{\bar{V}}_{\omega}(a, N_s; i)$.

Le problème de la permutation est résolu par l'utilisation de la méthode de corrélation des enveloppes vue dans la section 1.5.3. L'opérateur suivant est défini

$$\varepsilon \hat{\bar{V}}_{\omega}(a, N_s; i) = \frac{1}{2M} \sum_{N'_s = N_s - M}^{N_s + M} \sum_{j=1}^n |\hat{V}_{j,\omega}(a, N'_s; i)|. \quad (3.19)$$

où M est une constante positive et $\bar{V}_{j,\omega}(a, N_s; i)$ est l'élément j du vecteur $\bar{\bar{V}}_{\omega}(a, N_s; i)$.

Le produit intérieur de deux opérateurs est défini par

$$\varepsilon \bar{\bar{V}}_{\omega}(i) \cdot \varepsilon \bar{\bar{V}}_{\omega'}(j) = \sum_{N_s} \varepsilon \bar{\bar{V}}_{\omega}(N_s; i) \varepsilon \bar{\bar{V}}_{\omega'}(N_s; j) \quad (3.20)$$

et sa norme par

$$\left\| \varepsilon \bar{\bar{V}}_{\omega}(i) \right\| = \sqrt{\varepsilon \bar{\bar{V}}_{\omega}(i) \cdot \varepsilon \bar{\bar{V}}_{\omega}(i)} . \quad (3.21)$$

Voici la méthode de résolution de la permutation des fréquences étape par étape :

1. Trier ω par ordre de grandeur de la corrélation entre les éléments indépendants dans ω . Cette étape est validée en ordonnant les fréquences comme ci-dessous

$$sim(\omega) = \sum_{i \neq j} \frac{\varepsilon \bar{\bar{V}}_{\omega}(i) \varepsilon \bar{\bar{V}}_{\omega}(j)}{\left\| \varepsilon \bar{\bar{V}}_{\omega}(i) \right\| \left\| \varepsilon \bar{\bar{V}}_{\omega}(j) \right\|} \quad (3.22)$$

$$sim(\omega_1) \leq sim(\omega_2) \leq \dots \leq sim(\omega_N) \quad (3.23)$$

2. Pour ω_1 assigner $\bar{\bar{V}}_{\omega_1}(N_s; i)$ à $\bar{\bar{R}}_{\omega_1}(N_s; i)$ comme

$$\bar{\bar{R}}_{\omega_1}(N_s; i) = \bar{\bar{V}}_{\omega_1}(N_s; i), \quad i = 1, \dots, n_s \quad (3.24)$$

3. Pour chaque ω_d trouver une permutation $\sigma_{\omega}(i)$ qui maximise la corrélation en entre l'enveloppe de ω_d et l'enveloppe jointe (spectrogrammes divisés) de ω_1 à ω_{d-1} . Cette opération est effectuée en maximisant

$$\sum_{i=1}^{N_s} \varepsilon \hat{V}_{\omega_d}(\sigma_{\omega}(i)) \cdot \left(\sum_{j=1}^{d-1} \varepsilon \hat{R}_{\omega_j}(i) \right) \quad (3.25)$$

pour toute les permutations possibles σ_k avec $i = 1, \dots, n_s$.

4. Assigner la permutation appropriée à $\hat{\bar{R}}_{\omega_d}(N_s; i)$

$$\hat{\bar{R}}_{\omega_d}(N_s; i) = \hat{\bar{V}}_{\omega_d}(N_s; \sigma_{\omega}(i)) \quad (3.26)$$

Le résultat est la résolution de la permutation et l'obtention de spectrogrammes séparés définis par

$$\hat{\bar{R}}(\omega, N_s; i) = \hat{\bar{R}}_{\omega}(N_s; i). \quad (3.27)$$

Les coefficients reconstruits sont obtenus par la TFFGI des spectrogrammes séparés. Ils ont les mêmes dimensions que les observations et sont définies par

$$\bar{r}(a, n; i) = \frac{1}{2\pi} \frac{1}{W_f(n)} \sum_{N_s} \sum_{\omega} e^{-j\omega(n-N_s)} \hat{\bar{R}}(\omega, N_s; i), \quad i = 1, \dots, n_s \quad (3.28)$$

où i est l'indice des composantes indépendantes.

Pour chaque échelle les coefficients séparés permettent de reconstruire les sources séparées à conditions de vérifier la permutation entre chacune des échelles. La méthode utilisée est la même que celle employée précédemment pour la permutation au niveau des bandes de fréquences de la TFFG.

Le même opérateur est défini

$$\varepsilon \bar{r}_a(n; i) = \frac{1}{2M} \sum_{n'=n-M}^{n+M} \sum_{j=1}^{n_i} |r_{j,a}(n'; i)|. \quad (3.29)$$

où M est une constante positive et $r_{j,a}(n; i)$ signifie l'élément j du vecteur $\bar{r}_a(n; i)$.

Le produit intérieur de deux opérateurs est défini par

$$\varepsilon \bar{r}_a(i) \cdot \varepsilon \bar{r}_a(j) = \sum_n \varepsilon \bar{r}_a(n; i) \varepsilon \bar{r}_a(n; j) \quad (3.30)$$

et sa norme par

$$\|\varepsilon \bar{r}_a(i)\| = \sqrt{\varepsilon \bar{r}_a(i) \cdot \varepsilon \bar{r}_a(i)}. \quad (3.31)$$

Voici la méthode de résolution de la permutation des échelles étape par étape :

1. Trier a par ordre de grandeur de la corrélation entre les éléments indépendants dans a . Cette étape est accomplie en ordonnant les échelles comme ci-dessous

$$sim(a) = \sum_{i \neq j} \frac{\varepsilon \bar{r}_a(i) \varepsilon \bar{r}_a(j)}{\|\varepsilon \bar{r}_a(i)\| \|\varepsilon \bar{r}_a(j)\|} \quad (3.32)$$

$$sim(a_1) \leq sim(a_2) \leq \dots \leq sim(a_K) \quad (3.33)$$

2. Pour a_i assigner $\bar{r}_{a_i}(n; i)$ à $\bar{p}_{a_i}(n; i)$ comme

$$\bar{p}_{a_i}(n; i) = \bar{r}_{a_i}(n; i), \quad i = 1, \dots, n_s \quad (3.34)$$

3. Pour chaque a_d trouver une permutation $\sigma_a(i)$ qui maximise la corrélation entre l'enveloppe de a_d et l'enveloppe jointe des coefficients pour chaque échelle de a_i à a_{d-1} . Cette opération est effectuée en maximisant

$$\sum_{i=1}^{n_s} \varepsilon \bar{r}_{a_d}(\sigma_a(i)) \cdot \left(\sum_{j=1}^{d-1} \varepsilon \bar{p}_{a_j}(i) \right) \quad (3.35)$$

pour toute les permutations possibles σ_a avec $i = 1, \dots, n_s$.

4. Assigner la permutation appropriée à $\bar{p}_{a_d}(n; i)$

$$\bar{p}_{a_d}(n; i) = \bar{r}_{a_d}(n; \sigma_a(i)) \quad (3.36)$$

Le résultat est la résolution de la permutation et l'obtention des coefficients d'ondelettes séparés pour la synthèse.

$$\bar{p}(a, n; i) = \bar{p}_a(n; i). \quad (3.37)$$

3.3.4 Synthèse avec la transformée en ondelettes continues inverse

Pour la synthèse avec l'ondelette de Morlet une formule approximative est présente dans la littérature [55]. Celle-ci élimine l'intégration sur la translation, nécessitant seulement une intégration sur les échelles. La méthode trapézoïdale permet d'approcher cette intégration.

Pour les autres ondelettes la formule complète (3.39) est équivalente à la convolution de chaque échelle avec l'ondelette dilatée suivie d'une intégration finale sur l'ensemble des échelles.

La formule de synthèse approximative est définie par

$$\bar{p}(n;i) \approx K_{\psi} \int_0^{\infty} \bar{p}(a,n;i) \frac{da}{a^{3/2}} \quad (3.38)$$

3.3.5 Choix de la non-linéarité

Le choix de la non-linéarité dépend de la distribution statistique des sources recherchées. La vérification des distributions de toutes les composantes fréquentielles pour chacun des nœuds ne peut avoir lieu. Le nombre est beaucoup trop grand pour en faire la vérification.

La non-linéarité est choisie considérant une distribution laplacienne des composantes fréquentielles des coefficients de la TOC. La littérature fournit beaucoup d'information au sujet du choix des non-linéarités. La fonction sigmoïde est reconnue comme permettant d'obtenir une bonne solution pour les distributions laplaciennes [69]. Dans le cas de données complexes la contrainte imposée sur le choix de la non-linéarité est la préservation de la phase [23].

La non-linéarité choisie est définie par

$$g(\cdot) = \text{sign}(\cdot), \quad (3.39)$$

3.4 La séparation de sources avec la transformée en paquets d'ondelettes

Cette section résume rapidement la SS avec la TPO pour des fins de comparaison avec la SS utilisant la TOC. La SS avec l'utilisation de la TPO possède la même structure que la SS

avec la TOC. Les coefficients de chaque nœud sélectionnés sont séparés avec la méthode exposée dans la section 3.3. La méthode est identique à l'exception de la TO choisie. Le développement ne sera pas répété dans cette section car il a été fait à la section précédente. Il suffit de remplacer la TOC par la TPO.

La méthode est définie étape par étape comme suit :

1. Analyse de chaque mélange avec la TPO.
2. Pour chaque nœud de même position pour chaque arbre.
 - a. Analyse des coefficients avec la TFFG.
 - b. Pour chaque bande spectrale obtenir une matrice de séparation avec l'algorithme EASI
 - c. Résoudre les problèmes de dilatation et de la permutation en fréquence
 - d. Séparation des coefficients pour chaque nœud
 - e. Application de la TFFGI
2. Résoudre la permutation au niveau des nœuds
3. Effectuer la synthèse des coefficients pour obtenir les sources reconstruites

CHAPITRE 4

SIMULATIONS

4.1 Introduction

La première étape est l'application de l'algorithme de SS proposé au mélange instantané. Cette étape permet de vérifier le fonctionnement de l'algorithme avec les paramètres de simulation choisis. L'algorithme doit être en mesure de reconstruire les sources par rapport aux hypothèses faites sur les des distributions des sources.

La seconde étape est l'application de l'algorithme à des mélanges convolutifs provenant d'une matrice de mélanges constitués de filtres à RIF à phase minimale. Le séparateur est conçu pour permettre l'élimination des effets introduits par ces types de filtres.

4.2 Signaux sources

Les signaux sources proviennent de la base de données TIMIT [70]. Celle-ci contient des enregistrements de locuteurs effectuant la lecture d'une série de phrases pré définies dans un environnement anéchoïque. Pour la simulation une des sources est un locuteur et l'autre une locutrice provenant tout deux de la même région de dialecte.

4.3 Mélanges

Les mélanges utilisés pour l'évaluation des performances sont synthétiques. Les mélanges instantanés nécessitent le choix de coefficient pour l'obtention des mélanges. Ce choix est fait manuellement. Les mélanges convolutif nécessitent l'utilisation d'une boîte à outil MATLAB [71] pour l'obtention des fonctions de transfert nécessaire à l'obtention des mélanges.

4.3.1 Mélange instantané

La matrice de mélange instantané A est une matrice dont chaque coefficient correspond au poids des sources dans le mélange. La matrice de mélange instantané est choisie comme

$$A = \begin{bmatrix} 0.60 & 0.45 \\ 0.37 & 0.63 \end{bmatrix}. \quad (4.1)$$

Les mélanges sont obtenus conformément à l'équation 1.9 tel que

$$\begin{aligned} x_1(t) &= 0.60s_1(t) + 0.45s_2(t) \\ x_2(t) &= 0.37s_1(t) + 0.63s_2(t) \end{aligned} \quad (4.2)$$

4.3.2 Mélanges convolutifs

Les fonctions de transfert de la figure 4.3 sont obtenues par l'utilisation d'une boîte à outil MATLAB utilisant une méthode de calcul basée sur les images [71]. Ce modèle comporte les paramètres suivants : la dimension de la pièce, la position de la source, la position du capteur, le facteur d'absorption et le nombre de sources virtuelles. Une fonction de transfert à RIF a_{ij} est générée pour chaque configuration où i correspond à la source et j au capteur.

Par exemple a_{12} correspond à la réponse à l'impulsion entre la source 2 et le capteur 1.

La configuration des microphones M et des sources S à l'intérieur de l'espace virtuel est représenté aux figures 4.1 et 4.2. Les dimensions et les coordonnées sont en mètres. Le coefficient d'absorption est choisi comme $-1 \leq R \leq 1$ où -1 correspond au maximum et 1 au minimum d'absorption. Toutes les surfaces ont les mêmes propriétés dans ce modèle.

Les valeurs des paramètres sont les suivants :

- Coefficient d'absorption : $R = 0.5$
- Nombre de sources virtuelles : $n = 12, 25^3$ sources

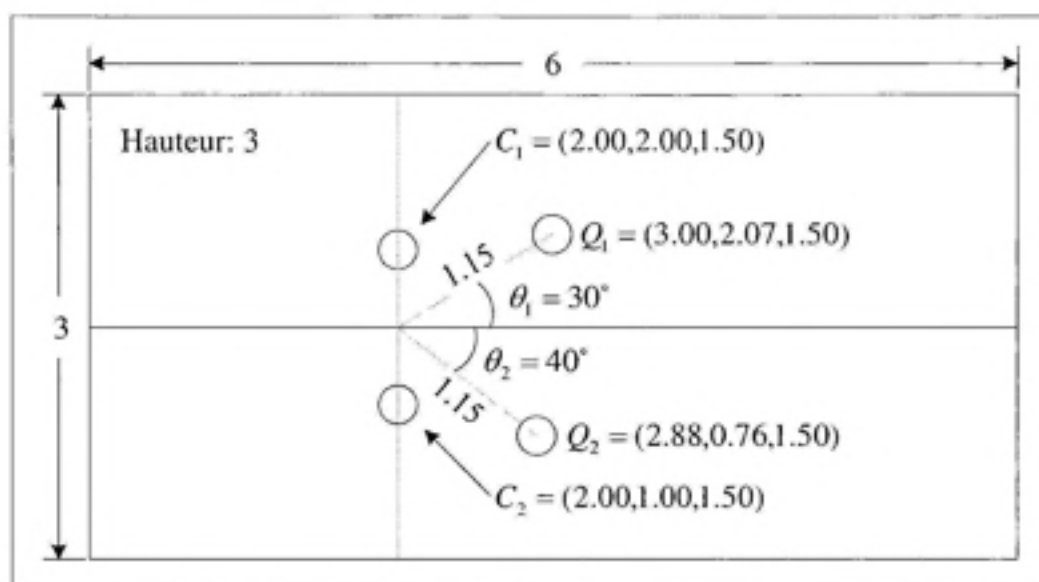


Figure 4.1 Configuration espace virtuel, vue de dessus.

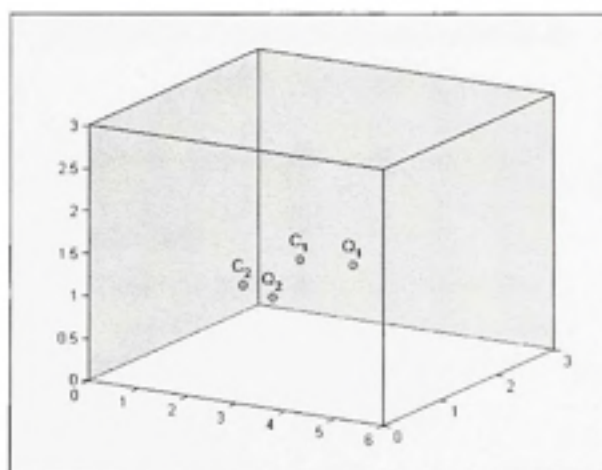


Figure 4.2 Configuration espace virtuel, vue isométrique.

Les fonctions de transfert obtenues sont représentées à la figure 4.3. La diaphonie au niveau des mélanges est ajustée pour éviter l'inversion d'une matrice singulière et pour permettre à l'algorithme de fonctionner.

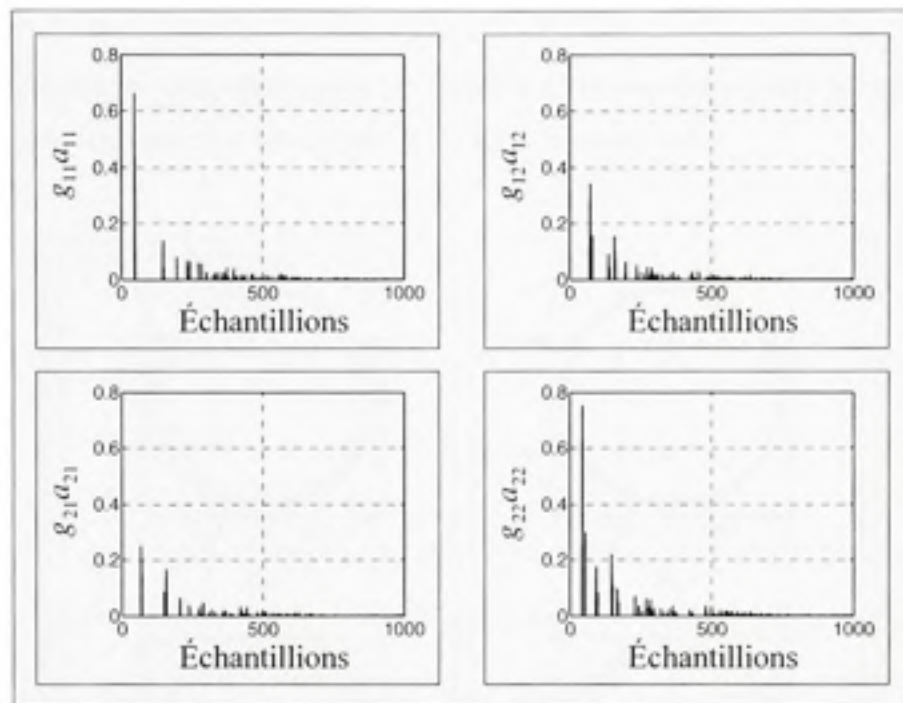


Figure 4.3 *Fonctions de transfert de la matrice de mélange.*

Les coefficients employés diminuent l'amplitude des fonctions de transfert normalisées directes et croisées. Leurs valeurs sont

$$G = \begin{bmatrix} 0.66 & 0.34 \\ 0.25 & 0.75 \end{bmatrix}. \quad (4.3)$$

Par multiplication de cette matrice de gain avec la matrice de mélange on obtient les mélanges finaux

$$\bar{x} = G \cdot A\bar{s}, \quad (4.4)$$

après développement

$$\begin{aligned} x_1(t) &= g_{11}a_{11} * s_1(t) + g_{12}a_{12} * s_2(t) \\ x_2(t) &= g_{21}a_{21} * s_1(t) + g_{22}a_{22} * s_2(t) \end{aligned} \quad (4.5)$$

Les zéros des fonctions sont représentés à la figure 4.4. Le modèle produit bien des fonctions à phase minimales car tous les zéros sont inclus dans le cercle unité.

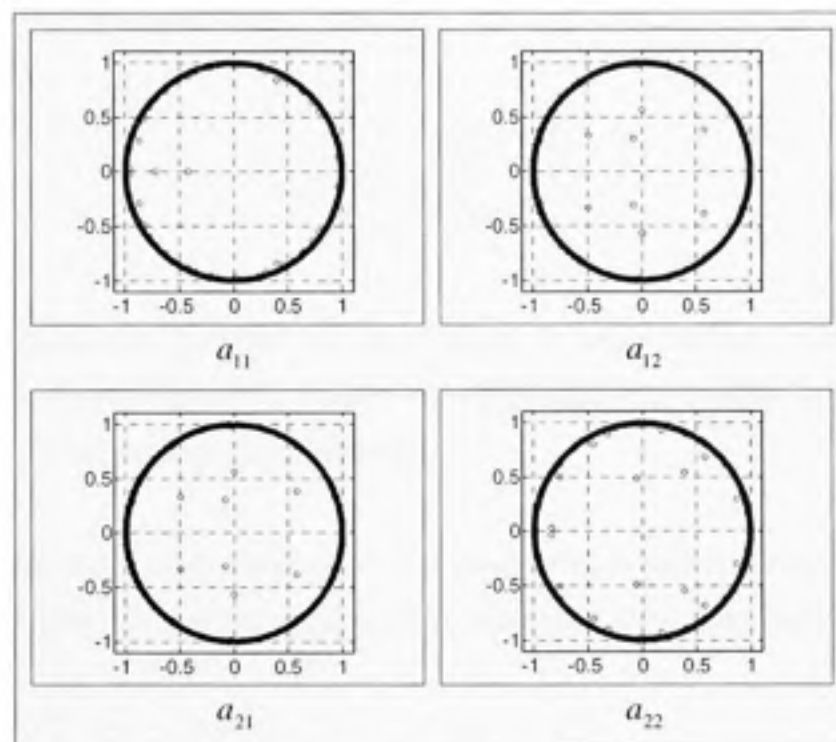


Figure 4.4 Zéros des fonctions de transfert de la matrice de mélange.

4.4 Évaluation des performances

Deux mesures sont employées pour évaluer les performances de l'algorithme : le niveau de distorsion et la qualité de la séparation [72]. Ces mesures sont possibles seulement si les sources sont connues. Dans le cas contraire des méthodes d'évaluation subjectives doivent être employées [49].

La distorsion est définie comme le niveau de différence entre la source d'origine et la source reconstruite. Elle est définie par

$$D_j = 10 \log \left(\frac{E\{(x_{j,s_j} - \alpha_j y_j)^2\}}{E\{(x_{j,s_j})^2\}} \right), \quad (4.6)$$

où $\alpha_j = E\{x_{j,s_j}^2\} / E\{x_j^2\}$ est un facteur d'échelle et j est l'indice de la source reconstruite à la sortie du système. La qualité de la séparation correspond au niveau de diaphonie à la sortie du système global. Elle est définie par

$$S_j = 10 \log \left(\frac{E\{(y_{j,s_j})^2\}}{E\{(\sum_{i \neq j} y_{j,s_i})^2\}} \right) \quad (4.7)$$

Toutes ces mesures sont possibles dû au principe de superposition. Les sources sont appliquées au système global et les signaux correspondant sont recueillis aux sorties ce qui permet de calculer les mesures de performances.

Avec la TOC la synthèse n'est pas parfaite. Les sources reconstruites doivent être comparées aux sources d'origine reconstruites à partir d'une analyse et d'une synthèse pour les échelles sélectionnées.

4.5 Paramètres de simulations

Les paramètres suivant sont fixes pour toutes les simulations :

- Pas : $\mu = 0.1$
- Non-linéarités : $g(\cdot) = \text{sign}(\cdot)$
- Nombre d'itérations : 75
- Fenêtre : Hamming

Le nombre de points de la TF et le déplacement de la fenêtre sont modulés pour obtenir une séparation optimale. Le nombre de point de la TF est défini par

$$N = 2^{k_N}, \quad k_N = 4, 5, \dots, 10 \quad (4.8)$$

et le déplacement de la fenêtre par

$$\Delta N = \frac{k_s N}{8}, \quad k_s = 1, 2, \dots, 6. \quad (4.9)$$

Les paramètres spécifiques aux deux algorithmes sont les suivant, pour la SS avec la TOC :

- Ondelette : Morlet $\omega_0 = 5$, $f_c = 13000$
- Paramètre de la loi logarithmique : $\mu_b = 8$
- Paramètre de la loi mu : $\mu_\sigma = 20$
- Nombre d'échelles : $K = 10$

pour la SS avec la TPO :

- Ondelette : Daubechies ordre 15

Le choix de l'ondelette de Morlet est justifié par son utilisation fréquente dans la littérature lors de l'analyse du signal de la parole [73-75]. De plus, ce choix permet l'utilisation de la formule de synthèse approximative (3.38). L'utilisation de l'ondelette de Daubechies est justifiée de la même façon. Cette ondelette est couramment utilisée lors de l'analyse du signal de la parole avec la TPO [76-78].

Les données sont générées par trois algorithmes de SS : avec la TOC, avec la TPO et sans prétraitement avec la TO. L'algorithme sans prétraitement avec la TO (en fréquence

seulement) est celui de la figure 4.5. Les données obtenues permettent de comparer la performance de l'algorithme en fonction du type de TO utilisé.

4.6 Résultats

Les tableaux qui suivent contiennent les données produites par la structure de simulation suivante :

- SS avec la TOC
 - Mélange instantané
 - Loi fréquences uniformes
 - Loi linéaire
 - Loi logarithmique
 - Loi mu
 - Mélange convolutif
 - Loi fréquences uniformes
 - Loi linéaire
 - Loi logarithmique
 - Loi mu
- SS avec la TPO
 - Mélange instantané
 - Niveau 1 à 3
 - Mélange convolutif
 - Niveau 1 à 3
- SS en fréquence seulement
 - Mélange instantané
 - Mélange convolutif

Chaque tableau contient un groupe de résultats. Chaque résultat correspond à une simulation pour une longueur de la fenêtre de la TFFG N et à un déplacement $\Delta N(\%)$.

Tableau 4.1

SS avec la TOC, mélange instantané, loi fréquences uniformes

	N								
	16	32	64	128	256	512	1024		
$\Delta N(\%)$	$S_{10} \text{ (dB)}$								
12.5	18.84	24.93	34.47	31.12	17.65	10.98	9.73		
25.0	32.15	35.00	33.32	32.25	20.49	11.45	14.93		
37.5	18.81	35.02	33.84	32.60	16.54	14.58	10.10		
50.0	18.81	35.46	35.88	32.45	16.27	12.46	12.55		
62.5	18.81	35.09	33.74	32.18	16.49	11.24	9.53		
75.0	18.82	34.98	34.11	32.18	16.42	12.11	10.12		
	$D_1 \text{ (dB)}$								
12.5	-17.55	-21.68	-31.16	-28.09	-13.07	-2.92	-3.31		
25.0	-32.10	-31.79	-31.40	-29.65	-16.93	-3.30	-7.25		
37.5	-17.53	-31.94	-31.22	-29.03	-12.30	-4.73	-3.40		
50.0	-17.53	-32.31	-32.07	-28.67	-12.09	-3.65	-4.23		
62.5	-17.53	-31.94	-31.58	-29.26	-12.26	-3.23	-3.15		
75.0	-17.54	-31.93	-31.47	-29.33	-12.16	-3.43	-3.46		
	$S_{20} \text{ (dB)}$								
12.5	18.45	27.18	31.96	29.00	18.30	7.64	7.17		
25.0	45.59	34.35	35.24	32.29	19.25	7.88	11.15		
37.5	18.46	34.75	33.27	29.91	16.12	9.11	7.41		
50.0	18.46	35.36	34.11	29.16	15.88	8.44	8.14		
62.5	18.46	34.60	34.84	31.28	16.09	7.91	7.36		
75.0	18.45	34.75	33.66	31.48	16.06	8.25	7.74		
	$D_2 \text{ (dB)}$								
12.5	-18.02	-20.51	-30.66	-27.62	-12.93	-5.91	-5.09		
25.0	-35.15	-31.70	-32.45	-30.10	-18.23	-6.28	-8.66		
37.5	-18.02	-31.95	-31.43	-28.57	-12.36	-8.07	-5.24		
50.0	-18.02	-32.38	-32.34	-27.95	-12.15	-6.70	-6.75		
62.5	-18.02	-31.89	-32.39	-29.42	-12.38	-6.17	-5.13		
75.0	-18.02	-31.94	-31.76	-29.54	-12.27	-6.52	-5.61		
	Temps de simulation (secondes)								
12.5	69	66	53	56	63	74	96		
25.0	76	74	59	62	69	81	103		
37.5	86	84	69	69	77	89	118		
50.0	102	106	80	79	89	103	135		
62.5	131	127	127	101	110	127	167		
75.0	185	182	181	148	152	174	220		
Échelles $a(k)$									
3.8	4.3	4.8	5.5	6.4	7.7	9.8	13.1	20.2	43.3
Diaphonie				Paramètres					
S_{10r}	S_{20r}	S_{10}	S_{20}	K	μ	n	Onl.		
3.86	3.26	2.50	4.62	10	0.10	75	Morlet		

Les résultats du tableau 4.1 montrent une bonne séparation des sources dans le cas du mélange instantané. Le niveau de séparation est élevé, le niveau de distorsion bas. Le choix des paramètres est justifié par la capacité de l'algorithme à séparer le mélange instantané.

Les performances diminuent avec l'augmentation du nombre de point de la TFFG. Les performances sont maximales pour les paramètres $N = 32$ et $\Delta T(\%) = 50$,

Tableau 4.2

SS avec la TOC, mélange instantané, loi linéaire

	N								
	16	32	64	128	256	512	1024		
$\Delta N(\%)$	$S_{10} \text{ (dB)}$								
12.5	29.81	33.30	34.60	24.31	17.63	14.65	12.45		
25.0	32.45	34.12	31.98	21.22	18.75	14.61	14.05		
37.5	29.71	33.86	33.76	30.72	18.77	12.46	11.73		
50.0	29.60	34.41	32.05	30.75	18.56	14.47	10.62		
62.5	29.65	34.16	34.06	29.26	18.68	14.25	10.47		
75.0	29.72	34.04	32.00	29.17	21.61	14.36	10.43		
	$D_1 \text{ (dB)}$								
12.5	-29.08	-31.15	-32.18	-20.21	-13.85	-8.48	-5.31		
25.0	-32.37	-30.86	-29.25	-14.55	-15.15	-6.85	-6.36		
37.5	-29.04	-30.71	-31.92	-26.62	-14.95	-5.28	-5.30		
50.0	-28.99	-31.16	-28.82	-26.43	-14.84	-7.07	-4.66		
62.5	-29.06	-30.96	-32.50	-25.07	-14.96	-6.69	-4.66		
75.0	-29.02	-30.87	-28.87	-24.86	-16.95	-6.86	-4.63		
	$S_{20} \text{ (dB)}$								
12.5	32.53	35.62	34.89	25.88	19.95	14.20	10.15		
25.0	44.13	34.94	34.82	19.60	21.47	12.50	11.13		
37.5	32.53	34.89	36.15	30.14	20.88	10.44	10.78		
50.0	32.53	35.30	33.18	29.66	20.77	12.83	9.48		
62.5	32.52	35.17	37.58	29.86	21.06	12.41	9.26		
75.0	32.52	35.00	33.42	29.56	22.45	12.58	9.34		
	$D_2 \text{ (dB)}$								
12.5	-31.14	-31.81	-33.10	-20.48	-14.13	-10.22	-7.35		
25.0	-36.01	-32.24	-30.49	-15.97	-15.16	-9.85	-8.60		
37.5	-31.15	-32.10	-33.61	-27.48	-14.91	-7.39	-7.59		
50.0	-31.15	-32.57	-29.69	-27.17	-14.82	-9.99	-6.18		
62.5	-31.21	-32.40	-34.53	-26.06	-14.96	-9.67	-5.96		
75.0	-31.14	-31.81	-33.10	-20.48	-14.13	-10.22	-7.35		
	Temps de simulation (secondes)								
12.5	73	71	59	63	68	78	100		
25.0	80	79	65	68	74	85	106		
37.5	91	89	74	74	82	94	122		
50.0	108	105	87	86	93	108	140		
62.5	137	132	134	107	115	133	167		
75.0	193	189	188	155	157	178	224		
Échelles $a(k)$									
3.8	8.2	12.6	17.0	21.4	25.8	30.2	34.6	38.9	43.3
Diaphonie				Paramètres					
S_{10r}	S_{20r}	S_{1l}	S_{2l}	K	μ	n	Ond.		
3.01	4.11	2.50	4.62	10	0.10	75	Morlet		

Les résultats du tableau 4.2 obtenus avec la loi linéaire sont sensiblement les mêmes que ceux obtenus avec la loi fréquences uniformes. Une augmentation significative des performances apparaît avec une fenêtre d'une longueur de 16 points. De plus une uniformisation des performances semble avoir lieu sur toute la plage simulée. Les performances sont maximales pour les paramètres $N = 32$ et $\Delta N(\%) = 50$.

Tableau 4.3

SS avec la TOC, mélange instantané, loi logarithmique

	N								
	16	32	64	128	256	512	1024		
ΔN (%)	S_{10} (dB)								
12.5	33.20	32.07	29.32	27.24	20.03	11.30	10.69		
25.0	33.37	31.90	29.28	26.75	25.28	17.66	12.44		
37.5	32.96	31.87	29.01	26.79	20.83	12.44	11.53		
50.0	33.03	32.28	29.12	26.97	26.20	17.19	8.28		
62.5	32.99	32.03	29.44	26.83	20.79	17.30	6.81		
75.0	33.13	32.08	29.16	26.98	25.63	17.18	6.99		
	D_1 (dB)								
12.5	-32.92	-29.19	-25.43	-22.33	-16.16	-7.05	-6.31		
25.0	-33.09	-28.89	-25.91	-22.17	-23.17	-12.08	-7.82		
37.5	-32.73	-28.87	-25.43	-22.26	-17.03	-7.24	-5.85		
50.0	-32.79	-29.19	-25.46	-22.28	-22.30	-12.10	-5.51		
62.5	-32.73	-29.03	-25.98	-22.26	-16.96	-11.86	-5.70		
75.0	-32.88	-29.07	-25.51	-22.34	-21.11	-11.58	-5.75		
	S_{20} (dB)								
12.5	41.12	35.51	32.24	29.08	23.02	14.23	12.64		
25.0	41.79	35.25	33.21	29.44	30.70	18.80	14.31		
37.5	41.78	35.25	32.69	29.34	24.61	13.97	12.13		
50.0	41.63	35.54	32.59	29.22	29.76	18.97	11.78		
62.5	41.33	35.52	33.31	29.37	24.40	18.66	12.93		
75.0	41.53	35.53	32.70	29.37	28.03	18.38	12.94		
	D_2 (dB)								
12.5	-36.82	-31.47	-27.00	-24.30	-17.76	-8.73	-7.72		
25.0	-37.00	-31.22	-27.58	-24.24	-26.13	-14.57	-9.44		
37.5	-36.83	-31.19	-27.14	-24.27	-18.33	-9.56	-8.21		
50.0	-36.85	-31.55	-27.11	-24.28	-24.73	-14.52	-6.13		
62.5	-36.73	-31.40	-27.61	-24.31	-18.34	-14.32	-5.42		
75.0	-36.90	-31.41	-27.15	-24.35	-23.70	-14.13	-5.57		
	Temps de simulation (secondes)								
12.5	77	74	63	66	71	82	103		
25.0	84	83	68	71	77	88	110		
37.5	95	92	78	78	85	98	125		
50.0	111	109	90	89	97	112	143		
62.5	139	135	137	110	118	136	171		
75.0	195	192	191	157	159	181	227		
	Échelles $a(k)$								
3.8	26.2	31.5	34.7	36.9	38.7	40.1	41.3	42.4	43.3
Diaphonie				Paramètres					
S_{10r}	S_{20r}	S_{D1}	S_{D2}	K	P_{μ}	μ	n	Ond.	
2.01	5.11	2.50	4.62	10	8.0	0.10	75	Morlet	

Une amélioration globale des performances est observée par rapport au deux lois précédentes. Les performances aux extrémités de la plage simulée sont améliorées par

l'utilisation de la loi logarithmique, surtout dans la région où la TFFG est la plus longue. Les performances sont maximales pour les paramètres $N = 16$ et $\Delta N(\%) = 25$.

Tableau 4.4

SS avec la TOC, mélange instantané, loi mu

	N								
	16	32	64	128	256	512	1024		
$\Delta N(\%)$	$S_{1s} (dB)$								
12.5	33.11	34.10	32.45	24.06	17.95	13.43	13.43		
25.0	32.27	34.20	32.05	30.75	18.75	13.48	15.16		
37.5	33.08	34.10	31.70	30.95	26.18	11.66	12.05		
50.0	32.81	34.54	32.25	30.78	18.42	13.70	11.44		
62.5	32.94	34.23	32.29	30.59	18.75	13.48	10.91		
75.0	32.88	34.22	32.17	30.69	18.62	13.64	11.28		
	$D_1 (dB)$								
12.5	-32.70	-31.69	-29.39	-20.27	-13.87	-5.17	-5.68		
25.0	-32.18	-31.32	-29.82	-27.53	-14.59	-5.41	-8.31		
37.5	-32.67	-31.28	-28.75	-27.01	-22.24	-4.28	-4.81		
50.0	-32.45	-31.82	-29.50	-26.89	-14.20	-5.67	-5.36		
62.5	-32.59	-31.42	-29.95	-27.25	-14.42	-5.45	-5.49		
75.0	-32.52	-31.49	-29.50	-27.32	-14.29	-5.58	-5.54		
	$S_{2s} (dB)$								
12.5	42.60	36.26	33.42	26.77	20.42	11.05	10.77		
25.0	45.15	35.74	35.80	33.12	21.13	11.18	13.61		
37.5	42.65	35.81	33.91	30.93	28.24	9.92	10.30		
50.0	42.68	36.79	34.35	30.71	20.64	11.55	10.47		
62.5	42.71	36.03	35.84	32.02	20.90	11.33	10.72		
75.0	42.70	36.24	34.46	32.07	20.84	11.45	10.61		
	$D_2 (dB)$								
12.5	-36.44	-33.33	-30.40	-20.81	-14.69	-8.76	-8.37		
25.0	-36.32	-32.96	-31.43	-29.15	-15.34	-8.82	-10.41		
37.5	-36.42	-32.95	-30.07	-28.15	-24.19	-7.24	-7.60		
50.0	-36.28	-33.66	-30.76	-27.93	-15.07	-9.11	-7.44		
62.5	-36.40	-33.14	-31.51	-28.66	-15.31	-8.92	-7.16		
75.0	-36.33	-33.23	-30.78	-28.70	-15.18	-9.03	-7.24		
	Temps de simulation (secondes)								
12.5	76	74	62	65	71	81	102		
25.0	83	81	66	70	77	87	109		
37.5	94	92	77	76	85	96	124		
50.0	110	107	88	88	97	110	141		
62.5	139	135	135	109	118	134	169		
75.0	194	191	189	155	158	180	226		
	Échelles $a(k)$								
3.8	19.0	25.8	30.3	33.6	36.2	38.4	40.2	41.9	43.3
	Diaphonie				Paramètres				
S_{1sr}	S_{2sr}	S_{D1}	S_{D2}	K	ν_a	μ	n	Ond.	
2.67	4.45	2.50	4.62	10	20.0	0.10	75	Morlet	

L'utilisation de la loi mu permet d'obtenir un maximum de séparation avec une TFFG d'une longueur de 16 points. Les résultats suivent ceux obtenus avec la loi logarithmique dû à la similitude entre ces dernières. Une augmentation très nette de la séparation apparaît au niveau de la deuxième source. Les performances sont maximales pour les paramètres $N = 16$ et $\Delta N(\%) = 12.5$.

Tableau 4.5

SS avec la TOC, mélange convolutif, loi fréquences uniformes

	N								
	16	32	64	128	256	512	1024		
$\Delta N(\%)$	$S_{10} \text{ (dB)}$								
12.5	3.83	3.40	3.61	7.35	5.79	9.55	7.19		
25.0	4.06	3.36	3.20	6.93	5.90	9.81	6.43		
37.5	3.97	3.35	3.35	6.83	6.09	9.54	5.35		
50.0	4.10	3.32	3.51	7.96	6.10	9.27	4.81		
62.5	4.01	3.34	3.48	7.77	6.11	9.47	6.40		
75.0	3.95	3.35	3.44	7.78	6.06	9.38	6.39		
	$D_1 \text{ (dB)}$								
12.5	-5.11	-5.02	-5.07	-6.40	-5.83	-4.47	-2.77		
25.0	-5.33	-4.97	-4.79	-6.09	-5.68	-4.60	-1.85		
37.5	-5.20	-4.97	-4.91	-6.09	-5.73	-3.73	-2.01		
50.0	-5.35	-4.94	-5.01	-6.45	-5.69	-4.79	-1.53		
62.5	-5.27	-4.96	-4.99	-6.52	-5.63	-4.75	-2.22		
75.0	-5.21	-4.96	-4.96	-6.40	-5.71	-4.44	-2.22		
	$S_{20} \text{ (dB)}$								
12.5	4.14	3.42	4.84	6.42	8.26	9.12	7.29		
25.0	3.87	3.42	4.60	5.63	7.94	9.97	7.44		
37.5	3.90	3.52	4.65	6.23	9.06	10.09	5.73		
50.0	3.72	3.18	4.73	5.59	8.93	9.77	4.41		
62.5	3.90	3.47	4.77	5.87	8.87	9.81	5.33		
75.0	4.01	3.29	4.75	5.81	8.81	10.11	5.54		
	$D_2 \text{ (dB)}$								
12.5	-5.40	-4.91	-5.82	-6.62	-5.67	-4.78	-4.85		
25.0	-5.25	-4.92	-5.64	-6.14	-5.60	-5.35	-3.08		
37.5	-5.25	-5.00	-5.68	-6.54	-5.59	-4.82	-3.67		
50.0	-5.15	-4.76	-5.73	-6.18	-5.55	-5.78	-3.33		
62.5	-5.26	-4.96	-5.76	-6.37	-5.43	-5.63	-3.84		
75.0	-5.33	-4.84	-5.74	-6.30	-5.61	-5.55	-4.12		
	Temps de simulation (secondes)								
12.5	86	91	77	70	78	89	111		
25.0	104	106	87	79	87	98	121		
37.5	126	119	107	91	104	111	141		
50.0	157	150	123	110	115	131	165		
62.5	178	205	197	139	146	165	203		
75.0	275	309	260	206	209	233	281		
	Échelles $a(k)$								
3.8	4.3	4.8	5.5	6.4	7.7	9.8	13.1	20.2	43.3
	Diaphonie			Paramètres					
S_{10r}	S_{20r}	S_{11}	S_{21}	K	μ	n	Ond.		
4.07	3.57	2.54	4.85	10	0.10	75	Morlet		

Le tableau 4.5 montre une augmentation des performances de séparation du mélange convolutif avec l'augmentation du nombre de point de la TFFG. Les performances sont maximales pour les paramètres $N = 512$ et $\Delta N(\%) = 50$.

Tableau 4.6

SS avec la TOC, mélange convolutif, loi linéaire

	<i>N</i>								
	16	32	64	128	256	512	1024		
$\Delta N(\%)$	<i>S_{1s} (dB)</i>								
12.5	2.88	5.44	5.44	6.67	7.63	7.78	6.68		
25.0	2.85	5.46	5.11	5.81	7.55	9.47	6.47		
37.5	2.87	5.47	4.16	5.67	7.61	10.65	6.82		
50.0	3.16	5.46	4.37	7.15	7.45	9.73	7.11		
62.5	2.86	5.49	4.31	6.94	7.64	10.10	8.24		
75.0	2.88	5.48	4.33	6.92	7.63	9.88	7.70		
	<i>D₁ (dB)</i>								
12.5	-4.60	-5.87	-6.28	-5.74	-6.11	-4.24	-4.01		
25.0	-4.58	-5.88	-6.09	-2.54	-5.31	-4.97	-3.60		
37.5	-4.60	-5.85	-5.53	-3.09	-6.06	-3.89	-3.22		
50.0	-4.79	-5.89	-5.69	-6.07	-6.41	-5.17	-3.20		
62.5	-4.59	-5.88	-5.64	-6.15	-6.14	-4.72	-4.07		
75.0	-4.60	-5.87	-5.66	-6.23	-6.46	-4.88	-4.03		
	<i>S_{2s} (dB)</i>								
12.5	4.82	5.47	6.54	4.89	7.60	7.89	7.54		
25.0	4.83	5.52	6.40	3.82	7.98	8.68	6.93		
37.5	4.82	5.53	5.09	3.92	7.76	9.68	8.17		
50.0	4.19	5.44	5.21	5.05	7.63	8.54	7.33		
62.5	4.80	5.51	5.18	5.05	7.61	8.99	7.20		
75.0	4.77	5.48	5.20	4.74	7.49	8.68	6.86		
	<i>D₂ (dB)</i>								
12.5	-5.78	-6.25	-6.89	-5.84	-6.34	-5.02	-4.67		
25.0	-5.79	-6.28	-6.84	-2.63	-5.94	-5.97	-4.53		
37.5	-5.75	-6.28	-6.43	-2.89	-6.50	-5.21	-3.79		
50.0	-5.44	-6.22	-6.43	-6.05	-6.68	-6.20	-4.20		
62.5	-5.76	-6.27	-6.43	-6.07	-6.38	-5.80	-4.93		
75.0	-5.73	-6.24	-6.43	-5.85	-6.59	-5.96	-5.18		
	<i>Temps de simulation (secondes)</i>								
12.5	86	85	73	76	83	95	117		
25.0	97	96	83	90	91	103	127		
37.5	111	110	98	95	103	117	146		
50.0	133	132	116	113	121	137	170		
62.5	172	171	173	143	152	171	209		
75.0	246	247	247	211	214	238	286		
	<i>Échelles a(k)</i>								
3.8	8.2	12.6	17.0	21.4	25.8	30.2	34.6	38.9	43.3
	<i>Diaphonie</i>				<i>Paramètres</i>				
<i>S_{far}</i>	<i>S_{2sr}</i>	<i>S_{2l}</i>	<i>S_{2l}</i>	<i>K</i>	μ	<i>n</i>	<i>Ond.</i>		
3.35	3.97	2.54	4.85	10	0.10	75	Morlet		

Les résultats du tableau 4.6 sont similaires sur toute la plage à ceux du tableau 4.5. Une augmentation de la séparation maximale pour la première source et une diminution pour la deuxième peut-être observée par rapport à la loi fréquences uniformes. Les performances sont maximales pour les paramètres $N = 512$ et $\Delta N(\%) = 50$.

Tableau 4.7

SS avec la TOC, mélange convolutif, loi logarithmique

	N								
	16	32	64	128	256	512	1024		
$\Delta N(\%)$	$S_{to} \text{ (dB)}$								
12.5	2.38	6.67	6.69	7.61	7.92	9.87	6.67		
25.0	2.33	6.61	6.63	7.86	8.79	10.58	7.53		
37.5	2.42	6.60	6.48	7.83	8.68	10.97	9.02		
50.0	2.44	6.50	6.58	7.92	8.67	10.09	8.90		
62.5	2.41	6.53	6.68	7.85	8.79	10.56	10.51		
75.0	2.36	6.52	6.63	7.85	8.80	10.65	8.89		
	$D_t \text{ (dB)}$								
12.5	-4.27	-6.44	-7.08	-7.18	-6.39	-5.70	-4.78		
25.0	-4.23	-6.73	-7.11	-7.62	-7.72	-6.50	-4.96		
37.5	-4.30	-6.70	-6.92	-7.66	-7.68	-5.62	-5.13		
50.0	-4.31	-6.52	-6.99	-7.42	-7.90	-5.94	-3.92		
62.5	-4.29	-6.53	-7.13	-7.59	-7.78	-5.42	-6.25		
75.0	-4.25	-6.47	-7.09	-7.61	-7.89	-5.76	-5.80		
	$S_{ss} \text{ (dB)}$								
12.5	7.11	8.05	8.69	8.81	10.34	12.22	10.81		
25.0	7.18	8.10	8.78	8.71	10.32	12.70	10.45		
37.5	7.06	8.12	8.67	8.40	10.32	13.03	13.31		
50.0	6.80	8.15	8.70	8.67	10.31	12.58	12.14		
62.5	7.07	8.08	8.67	8.76	10.55	12.94	13.68		
75.0	7.13	8.16	8.73	8.73	10.48	12.92	12.74		
	$D_s \text{ (dB)}$								
12.5	-7.40	-8.35	-8.80	-8.26	-8.28	-8.49	-7.56		
25.0	-7.40	-8.44	-8.75	-8.20	-9.44	-9.45	-7.69		
37.5	-7.35	-8.45	-8.73	-8.05	-9.40	-8.98	-7.62		
50.0	-7.32	-8.44	-8.74	-8.16	-9.43	-9.02	-5.64		
62.5	-7.36	-8.40	-8.77	-8.23	-9.55	-8.35	-8.74		
75.0	-7.39	-8.41	-8.81	-8.20	-9.52	-8.62	-8.52		
	Temps de simulation (secondes)								
12.5	89	88	76	77	85	96	119		
25.0	100	99	85	85	93	106	129		
37.5	113	113	101	97	106	119	149		
50.0	136	135	117	115	123	139	173		
62.5	174	174	175	146	154	174	211		
75.0	249	250	250	214	217	240	288		
	Échelles a(k)								
3.8	26.2	31.5	34.7	36.9	38.7	40.1	41.3	42.4	43.3
	Diaphonie				Paramètres				
S_{liv}	S_{2lv}	S_{ll}	S_{2l}	K	μ_b	μ	n	Ond.	
3.44	6.40	2.54	4.85	10	8.0	0.10	75	Morlet	

L'utilisation de la loi logarithme accroît les performances de la séparation comme dans le cas du mélange instantané. Le niveau de performance optimal est toujours situé dans la colonne où une TFFG de 512 est utilisée. Le niveau de séparation augmente et le niveau de distorsion

diminue comparativement à l'utilisation de la loi fréquence et échelles uniformes. Les performances sont maximales pour les paramètres $N = 512$ et $\Delta N(\%) = 50$.

Tableau 4.8

SS avec la TOC, mélange convolutif, loi mu

	N								
	16	32	64	128	256	512	1024		
$\Delta N(\%)$	$S_{10} \text{ (dB)}$								
12.5	3.12	5.20	4.22	4.85	6.66	8.96	5.79		
25.0	3.17	5.17	4.04	4.69	7.99	9.79	6.77		
37.5	3.17	5.17	4.05	4.49	7.49	10.12	6.75		
50.0	3.33	5.14	4.15	4.76	8.10	9.26	7.68		
62.5	3.18	5.16	4.12	4.96	8.10	10.26	8.43		
75.0	3.13	5.14	4.12	4.88	8.18	9.69	6.92		
	$D_1 \text{ (dB)}$								
12.5	-4.77	-5.76	-5.65	-5.56	-5.54	-4.45	-2.98		
25.0	-4.80	-5.87	-5.54	-5.60	-6.09	-5.46	-3.40		
37.5	-4.81	-5.86	-5.55	-5.14	-6.17	-5.14	-4.29		
50.0	-4.93	-5.74	-5.59	-5.44	-6.53	-5.30	-3.82		
62.5	-4.82	-5.73	-5.59	-5.64	-6.30	-5.17	-4.01		
75.0	-4.78	-5.72	-5.58	-5.59	-6.44	-5.15	-3.55		
	$S_{20} \text{ (dB)}$								
12.5	5.72	6.35	6.21	6.26	9.64	10.15	8.62		
25.0	5.68	6.22	6.20	5.28	7.61	11.05	8.32		
37.5	5.62	6.21	6.24	5.78	7.39	11.10	9.10		
50.0	5.35	6.22	6.31	5.38	7.53	10.57	9.49		
62.5	5.64	6.28	6.34	5.88	7.46	11.14	8.44		
75.0	5.72	6.22	6.29	5.64	7.43	11.03	7.66		
	$D_2 \text{ (dB)}$								
12.5	-6.38	-6.85	-7.19	-6.94	-6.58	-6.67	-5.22		
25.0	-6.37	-6.88	-7.14	-6.21	-7.19	-7.64	-5.70		
37.5	-6.27	-6.89	-7.24	-6.41	-7.16	-7.61	-5.69		
50.0	-6.20	-6.77	-7.20	-6.18	-7.34	-7.37	-5.23		
62.5	-6.32	-6.81	-7.26	-6.71	-7.24	-7.26	-6.17		
75.0	-6.36	-6.77	-7.20	-6.50	-7.25	-7.35	-6.13		
	Temps de simulation (secondes)								
12.5	88	87	75	77	84	96	120		
25.0	99	98	84	84	92	105	129		
37.5	112	111	100	97	104	118	149		
50.0	134	134	116	115	122	137	172		
62.5	173	173	174	145	153	173	211		
75.0	248	248	249	213	216	239	289		
Échelles $a(k)$									
3.8	19.0	25.8	30.3	33.6	36.2	38.4	40.2	41.9	43.3
Diaphonie				Paramètres					
S_{10r}	S_{20r}	S_R	S_N	K	μ_a	μ	n	Ond.	
3.45	5.08	2.54	4.85	10	20.0	0.10	75	Morlet	

Les résultats du tableau 4.8 sont sensiblement les mêmes que ceux obtenus avec la loi fréquence linéaire et fréquences uniformes. Le niveau de performance optimal est toujours dans la colonne où une TFFG de 512 points est employé. Les performances sont maximales pour les paramètres $N = 512$ et $\Delta N(\%) = 50$.

Tableau 4.9

SS avec la TPO, mélange instantané, analyse niveau 1

	<i>N</i>						
	16	32	64	128	256	512	1024
$\Delta N(\%)$	<i>S₁₀ (dB)</i>						
12.5	37.94	35.04	7.97	12.57	8.72	7.02	15.80
25.0	39.59	13.65	7.88	18.39	7.55	7.75	8.07
37.5	39.24	36.58	7.88	18.29	8.49	9.01	8.49
50.0	39.39	38.65	7.89	18.26	8.60	8.82	6.14
62.5	38.91	38.12	7.90	18.31	8.70	8.49	7.80
75.0	38.64	38.36	7.91	18.33	8.70	8.30	7.72
	<i>D₁ (dB)</i>						
12.5	3.83	3.82	3.29	3.45	3.11	2.82	3.71
25.0	3.80	3.52	3.30	3.59	2.89	3.02	3.24
37.5	3.80	3.80	3.32	3.63	3.03	3.18	3.14
50.0	3.82	3.80	3.31	3.63	3.07	3.21	2.63
62.5	3.81	3.80	3.30	3.62	3.09	3.10	3.01
75.0	3.82	3.80	3.30	3.62	3.09	3.10	3.00
	<i>S₂₀ (dB)</i>						
12.5	35.23	37.31	9.39	11.24	11.87	8.36	11.42
25.0	37.59	15.71	9.46	21.49	7.17	9.86	2.89
37.5	35.72	34.61	9.52	21.96	10.89	10.12	6.02
50.0	37.58	36.33	9.51	21.71	11.39	12.20	7.24
62.5	37.30	35.59	9.51	21.75	11.51	9.69	6.87
75.0	37.41	35.66	9.51	21.78	11.52	10.84	6.95
	<i>D₂ (dB)</i>						
12.5	5.13	5.16	4.54	4.95	5.06	4.76	4.80
25.0	5.19	5.10	4.51	5.29	4.57	4.79	3.42
37.5	5.18	5.18	4.50	5.25	5.03	4.86	4.29
50.0	5.16	5.19	4.52	5.25	5.00	4.94	4.62
62.5	5.16	5.19	4.52	5.27	5.03	4.81	4.48
75.0	5.16	5.19	4.52	5.26	5.03	4.88	4.50
	<i>Temps de simulation (secondes)</i>						
12.5	13	12	12	13	14	16	20
25.0	14	12	13	13	14	16	21
37.5	16	16	14	14	16	18	22
50.0	18	18	16	16	17	20	24
62.5	21	21	19	19	21	23	28
75.0	28	29	30	25	26	30	36
<i>Diaphonie</i>				<i>Paramètres</i>			
<i>S₁₁</i>	<i>S₂₁</i>	<i>Niveau</i>		μ	<i>n</i>	<i>Ond.</i>	
2.54	4.85	1		0.10	75	Db15	

Les performances exposées dans le tableau 4.9 diminuent rapidement avec l'augmentation de la longueur de la TFFG comparativement à l'utilisation de la TOC pour effectuer la SS. Le niveau de distorsion pour le niveau 1 avec la TPO est dû à une erreur au niveau de la permutation des nœuds lors de la reconstruction. Les performances sont maximales pour les paramètres $N = 512$ et $\Delta N(\%) = 50$.

Tableau 4.10

SS avec la TPO, mélange instantané, analyse niveau 2

$\Delta N(\%)$	N						
	16	32	64	128	256	512	1024
	S_{μ} (dB)						
12.5	35.44	35.24	5.58	9.84	10.12	10.70	3.32
25.0	37.55	36.23	5.66	4.43	9.41	7.31	4.21
37.5	37.04	34.68	5.50	5.17	11.52	9.06	6.57
50.0	39.17	36.97	5.62	4.95	8.54	10.39	6.53
62.5	38.75	36.33	5.64	5.00	9.39	8.02	6.19
75.0	38.96	36.20	5.63	5.07	9.07	8.48	5.97
	D_1 (dB)						
12.5	3.46	3.37	2.14	2.87	2.85	3.00	1.61
25.0	3.44	3.42	2.16	1.83	2.93	2.50	1.42
37.5	3.43	3.43	2.14	1.99	3.03	2.67	2.23
50.0	3.44	3.44	2.17	1.95	2.67	2.92	2.27
62.5	3.43	3.43	2.17	1.98	2.83	2.62	2.16
75.0	3.43	3.43	2.16	1.99	2.79	2.72	2.14
	S_{20} (dB)						
12.5	37.35	31.83	4.08	8.08	7.33	7.66	0.00
25.0	35.20	32.93	4.15	3.12	12.94	4.44	2.13
37.5	34.95	34.41	4.19	5.01	14.98	5.68	5.58
50.0	36.91	35.63	4.15	4.95	10.58	9.40	2.87
62.5	36.01	33.65	4.15	5.05	12.46	7.01	2.53
75.0	35.98	33.75	4.16	5.08	11.70	7.68	2.59
	D_2 (dB)						
12.5	4.63	4.75	3.23	4.13	3.97	3.96	1.06
25.0	4.67	4.69	3.25	2.82	4.39	3.38	2.47
37.5	4.66	4.69	3.20	3.50	4.56	3.75	3.70
50.0	4.67	4.68	3.22	3.41	4.38	4.20	2.93
62.5	4.67	4.69	3.23	3.47	4.46	3.93	2.79
75.0	4.67	4.69	3.24	3.50	4.40	4.02	2.77
	Temps de simulation (secondes)						
12.5	12	12	13	14	16	20	27
25.0	13	13	14	15	16	21	27
37.5	16	14	15	16	18	22	29
50.0	18	16	17	17	20	23	31
62.5	22	19	19	21	23	27	35
75.0	29	29	25	26	28	34	41
Diaphonie				Paramètres			
S_{μ}	S_{20}	Niveau	μ	n	Ond.		
2.54	4.85	2	0.10	75	Db15		

Les résultats du tableau 4.10 sont similaires à ceux du tableau 4.9. Le passage du niveau 1 au niveau 2 de décomposition ne semble pas avoir un impact particulier sur les performances. La distorsion est toujours positive dû à une erreur de permutation des nœuds. Les performances sont maximales pour les paramètres $N = 512$ et $\Delta N(\%) = 50$.

Tableau 4.11

SS avec la TPO, mélange instantané, analyse niveau 3

$\Delta N(\%)$	N						
	16	32	64	128	256	512	1024
$S_{10} (dB)$							
12.5	21.17	5.71	7.46	6.35	11.36	8.10	3.45
25.0	23.81	5.79	7.30	9.46	7.26	6.80	4.82
37.5	23.80	5.61	4.50	9.11	8.40	7.72	6.45
50.0	21.07	5.77	7.29	9.13	7.85	7.34	5.08
62.5	21.08	5.79	7.21	9.27	7.16	8.80	6.56
75.0	21.06	5.78	7.04	8.01	7.67	8.51	7.31
$D_1 (dB)$							
12.5	2.92	1.63	1.95	1.87	-5.67	-5.89	-1.72
25.0	2.95	1.65	1.92	-5.54	-4.94	-5.26	-2.13
37.5	2.98	1.59	1.22	-6.24	2.20	-3.67	-3.66
50.0	3.01	1.67	1.88	2.28	2.13	-4.87	-2.12
62.5	2.99	1.66	1.88	2.33	1.98	-5.51	-3.32
75.0	2.99	1.66	1.82	2.13	2.05	-5.24	-3.83
$S_{20} (dB)$							
12.5	19.38	3.99	6.86	5.11	10.81	11.06	6.10
25.0	23.04	4.05	6.78	11.44	11.42	11.56	6.93
37.5	23.29	4.02	3.75	13.19	5.50	9.22	9.42
50.0	19.96	4.09	6.40	6.16	6.04	11.17	7.44
62.5	19.89	4.06	6.44	7.97	4.46	11.70	8.54
75.0	19.89	4.08	6.09	7.53	4.99	11.40	9.12
$D_2 (dB)$							
12.5	4.60	2.94	3.79	3.19	-7.34	-6.27	-2.79
25.0	4.50	2.96	3.69	-6.17	-4.87	-5.70	-3.34
37.5	4.48	2.90	2.83	-6.75	3.37	-4.88	-5.44
50.0	4.49	2.94	3.67	3.57	3.50	-4.83	-3.52
62.5	4.51	2.96	3.69	3.87	3.14	-6.03	-4.68
75.0	4.51	2.96	3.66	3.82	3.33	-5.85	-5.29
<i>Temps de simulation (secondes)</i>							
12.5	90	89	77	79	86	96	119
25.0	101	101	85	86	95	106	129
37.5	115	115	118	99	106	119	148
50.0	138	137	141	117	123	138	172
62.5	176	176	180	145	155	175	210
75.0	252	254	257	261	218	241	287
<i>Diaphonie</i>				<i>Paramètres</i>			
S_{μ}	S_N	Niveau		μ	n	Ond.	
2.54	4.85	3		0.10	75	Db15	

Avec une décomposition de niveau 3 une dégradation des performances commence à apparaître par rapport aux niveaux 1 et 2. La SS est arrêtée à ce niveau dû à une diminution évidente des performances avec une décomposition supérieure au niveau 3. Les performances sont maximales pour les paramètres $N = 512$ et $\Delta N(\%) = 50$.

Tableau 4.12

SS avec la TPO, mélange convolutif, analyse niveau 1

	<i>N</i>						
	16	32	64	128	256	512	1024
$\Delta N(\%)$	<i>S_{1s}</i> (dB)						
12.5	2.20	4.53	6.10	6.43	3.04	3.18	5.07
25.0	2.48	4.14	6.57	6.67	2.95	5.84	2.20
37.5	2.68	4.13	7.01	4.44	2.80	6.83	2.61
50.0	2.61	6.36	4.88	4.60	2.71	6.73	1.92
62.5	2.53	4.25	5.73	6.46	2.99	2.88	1.47
75.0	2.48	4.36	5.51	6.45	3.13	6.83	2.18
	<i>D₁</i> (dB)						
12.5	-4.24	-5.54	4.32	-5.08	2.63	2.86	-3.97
25.0	-4.42	6.47	4.70	-5.10	2.69	-3.49	2.87
37.5	-4.54	-4.98	3.99	3.31	2.94	-4.52	3.09
50.0	-4.50	7.45	3.71	3.37	3.05	-4.50	3.08
62.5	-4.45	-5.38	4.24	-5.17	2.75	3.17	2.88
75.0	-4.41	-5.45	4.07	-5.17	2.75	-4.58	3.08
	<i>S_{2s}</i> (dB)						
12.5	-6.06	-7.49	6.16	-4.03	3.89	3.59	-3.97
25.0	-6.26	7.74	6.59	-3.95	3.80	-4.66	3.04
37.5	-6.35	-7.17	6.41	3.89	4.25	-4.89	3.75
50.0	-6.28	8.04	6.14	4.06	4.88	-5.02	3.48
62.5	-6.22	-7.45	6.54	-3.74	4.11	3.85	3.67
75.0	-6.30	-7.49	6.46	-3.75	3.91	-4.92	3.40
	<i>Temps de simulation (secondes)</i>						
12.5	8	7	7	7	9	11	15
25.0	9	7	8	8	9	11	16
37.5	11	9	9	10	10	13	17
50.0	13	11	10	11	12	15	19
62.5	16	16	13	14	16	18	23
75.0	24	24	20	20	22	25	31
<i>Diaphonie</i>		<i>Paramètres</i>					
<i>S_{1f}</i>	<i>S_{2f}</i>	<i>Niveau</i>	μ	<i>n</i>	<i>Ond.</i>		
2.54	4.85	1	0.10	75	Db15		

Les performances au niveau 1 sont maximales avec une TFFG de 64 points. Encore un fois le niveau de distorsion est positif signifiant un erreur de reconstruction des sources. Les performances sont maximales pour les paramètres $N = 512$ et $\Delta N(\%) = 50$.

Tableau 4.13

SS avec la TPO, mélange convolutif, analyse niveau 2

	N						
	16	32	64	128	256	512	1024
$\Delta N(\%)$	$S_{10} (dB)$						
12.5	4.48	2.35	6.44	5.26	6.54	2.71	5.04
25.0	4.00	5.06	6.56	4.99	5.82	2.04	5.03
37.5	4.18	2.43	6.50	4.89	6.85	5.52	4.98
50.0	4.40	2.26	6.64	4.90	6.70	2.04	5.30
62.5	4.18	2.57	6.62	4.88	6.98	2.72	4.74
75.0	4.34	2.54	6.62	5.00	6.93	4.03	5.41
	$D_1 (dB)$						
12.5	-5.52	-2.51	-4.91	-3.37	-3.94	1.88	-2.88
25.0	-5.22	-5.47	-5.05	-3.56	-3.70	1.98	-2.81
37.5	-5.38	-2.60	-5.01	-3.50	-4.03	-3.29	-3.56
50.0	-5.51	-2.48	-5.10	-3.59	-4.21	1.95	-2.37
62.5	-5.36	-2.68	-5.02	-3.59	-4.25	1.94	-2.82
75.0	-5.45	-2.63	-5.04	-3.61	-4.18	-2.96	-2.80
	$S_{20} (dB)$						
12.5	7.08	5.32	7.00	7.79	7.42	3.15	6.92
25.0	6.86	6.68	7.59	7.70	6.31	2.09	6.12
37.5	6.87	5.26	7.59	5.98	7.77	5.40	6.43
50.0	6.89	4.75	6.95	6.26	7.60	2.38	7.41
62.5	6.94	5.21	7.07	6.10	7.26	2.45	6.64
75.0	7.01	4.99	7.05	7.60	7.25	5.30	6.75
	$D_2 (dB)$						
12.5	-7.46	-3.30	-3.48	-5.03	-4.94	2.62	-3.97
25.0	-7.37	-7.02	-3.87	-5.10	-4.66	2.27	-3.56
37.5	-7.37	-3.48	-4.06	-5.02	-4.78	-3.77	-3.95
50.0	-7.37	-3.22	-3.69	-5.16	-5.01	1.99	-3.32
62.5	-7.41	-3.39	-3.60	-5.12	-5.07	2.14	-3.86
75.0	-7.46	-3.31	-3.63	-5.20	-4.93	-3.36	-3.77
	<i>Temps de simulation (secondes)</i>						
12.5	7	7	8	9	11	15	21
25.0	8	8	9	10	11	15	22
37.5	9	9	10	11	13	16	24
50.0	11	11	11	12	15	18	26
62.5	17	13	14	15	17	22	29
75.0	23	21	20	21	23	28	36
<i>Diaphonie</i>				<i>Paramètres</i>			
S_{11}	S_{21}	<i>Niveau</i>		μ	n	<i>Ond.</i>	
2.54	4.85	2		0.10	75	Db15	

Les résultats obtenus au niveau 2 sont beaucoup plus significatif que ceux du niveau 1. Les valeurs de la qualité de la séparation sont positives et les valeurs de la distorsion négatives sur la majeure partie de la plage. Ce sont les meilleurs résultats obtenus avec les paramètres choisis et l'utilisation de la TPO. Les performances sont maximales pour les paramètres $N = 512$ et $\Delta N(\%) = 50$.

Tableau 4.14

SS avec la TPO, mélange convolutif, analyse niveau 3

	<i>N</i>						
	16	32	64	128	256	512	1024
$\Delta N(\%)$	<i>S_{ta} (dB)</i>						
12.5	5.22	4.46	6.24	4.53	6.62	3.35	3.09
25.0	-0.81	3.20	5.52	4.66	4.77	5.00	3.15
37.5	2.13	3.46	5.66	6.02	4.17	5.38	3.73
50.0	3.86	3.62	5.67	6.31	3.42	4.18	2.54
62.5	2.23	3.46	5.89	5.66	4.96	2.89	3.43
75.0	2.15	3.76	5.46	6.01	3.60	4.53	3.04
	<i>D₁ (dB)</i>						
12.5	-4.45	-2.53	-3.76	-2.20	-2.60	-2.05	-1.89
25.0	0.57	-0.94	-3.47	-2.37	-2.78	-2.08	-1.99
37.5	-2.14	-0.29	-3.32	-2.74	-2.43	-1.67	-2.10
50.0	-2.24	-0.26	-3.53	-2.97	-2.27	-2.11	-1.39
62.5	-2.24	-0.30	-3.59	-2.09	-2.04	-1.31	-1.46
75.0	-2.18	-0.01	-3.02	-2.64	-2.33	-1.57	-1.14
	<i>S_{2a} (dB)</i>						
12.5	6.13	7.09	7.56	5.74	5.96	6.07	4.79
25.0	-0.97	5.85	7.63	6.19	5.87	6.21	5.39
37.5	5.71	6.36	7.71	7.68	5.94	6.07	5.98
50.0	6.95	6.25	7.77	7.09	5.25	6.27	4.66
62.5	5.58	6.31	7.73	7.60	5.69	6.23	5.29
75.0	5.40	5.91	7.94	7.65	5.41	5.94	5.68
	<i>D₂ (dB)</i>						
12.5	-3.92	-2.68	-5.11	-3.74	-3.78	-3.62	-2.88
25.0	1.10	-1.97	-4.59	-3.50	-2.96	-3.42	-2.60
37.5	-3.41	-1.69	-4.51	-3.73	-3.23	-2.88	-2.91
50.0	-4.42	-1.61	-4.80	-4.73	-3.26	-3.24	-2.60
62.5	-3.34	-1.65	-5.05	-3.51	-3.44	-2.81	-2.48
75.0	-3.27	-1.27	-4.48	-4.03	-3.35	-2.74	-2.37
	<i>Temps de simulation (secondes)</i>						
12.5	8	8	9	11	15	22	35
25.0	9	9	10	12	16	22	35
37.5	10	10	11	13	17	24	36
50.0	11	12	13	15	18	26	39
62.5	14	14	16	18	21	28	42
75.0	20	20	21	23	27	34	48
<i>Diaphonie</i>			<i>Paramètres</i>				
<i>S_{II}</i>	<i>S_{NI}</i>	<i>Niveau</i>	μ	<i>n</i>	<i>Ond.</i>		
2.54	4.85	3	0.10	75	Db15		

Comme dans le cas du mélange instantané une dégradation des performances est observable à partir du troisième niveau de décomposition. Les performances optimales obtenues sont inférieures par rapport au niveau 2. Les simulations ne sont pas poursuivies au-delà de ce niveau pour cette raison. Les résultats optimaux sont localisés dans la colonne où la TFFG est de 64 points.

Tableau 4.15

SS en fréquence, mélange instantané

	<i>N</i>						
	16	32	64	128	256	512	1024
$\Delta N(\%)$	<i>S₁₀ (dB)</i>						
12.5	8.63	15.55	12.63	15.94	17.82	10.44	9.46
25.0	8.61	15.81	12.52	11.49	18.00	10.76	10.97
37.5	8.64	15.64	12.47	11.48	18.15	14.01	17.14
50.0	8.65	15.76	12.52	16.19	12.31	10.81	12.44
62.5	8.64	15.75	12.50	16.20	18.09	10.83	15.06
75.0	8.63	15.75	12.51	16.19	18.10	10.83	15.59
	<i>D₁ (dB)</i>						
12.5	-4.03	-10.04	-6.12	-9.45	-14.03	-4.06	-6.12
25.0	-4.04	-10.16	-6.16	-6.71	-14.13	-5.57	-5.03
37.5	-4.05	-10.14	-6.11	-6.71	-14.13	-8.55	-10.33
50.0	-4.06	-10.16	-6.12	-9.48	-9.14	-5.67	-7.43
62.5	-4.05	-10.16	-6.14	-9.47	-14.15	-5.64	-8.25
75.0	-4.05	-10.15	-6.13	-9.47	-14.15	-5.64	-8.64
	<i>S₂₀ (dB)</i>						
12.5	9.73	15.48	11.58	14.85	20.93	10.02	12.76
25.0	9.76	15.68	11.82	12.03	21.20	11.34	10.74
37.5	9.77	15.71	11.70	12.05	20.86	14.74	15.82
50.0	9.77	15.68	11.71	14.78	16.07	11.69	13.57
62.5	9.76	15.64	11.78	14.75	21.07	11.62	14.01
75.0	9.76	15.62	11.75	14.77	21.06	11.64	14.32
	<i>D₂ (dB)</i>						
12.5	-5.42	-11.30	-8.27	-11.17	-14.78	-6.57	-7.22
25.0	-5.41	-11.44	-8.34	-7.39	-14.89	-6.97	-7.08
37.5	-5.42	-11.41	-8.29	-7.38	-14.80	-10.24	-12.65
50.0	-5.42	-11.42	-8.30	-11.18	-9.16	-7.10	-9.13
62.5	-5.43	-11.40	-8.33	-11.18	-14.85	-7.07	-10.81
75.0	-5.41	-11.39	-8.32	-11.18	-14.84	-7.08	-11.24
	<i>Temps de simulation (secondes)</i>						
12.5	13	12	11	11	12	13	16
25.0	14	14	12	12	13	14	17
37.5	15	15	15	13	14	16	19
50.0	17	17	17	15	16	18	21
62.5	21	21	21	18	19	21	25
75.0	29	29	29	30	25	28	32
<i>Diaphonie</i>				<i>Paramètres</i>			
<i>S₁₁</i>		<i>S₂₁</i>		μ		<i>n</i>	
2.50		4.62		0.10		75	

Les performances obtenues avec l'algorithme de SS en fréquence dans le cas du mélange instantané sont inférieures à celles obtenus avec l'utilisation de la TPO et de la TOC. Le niveau de séparation est beaucoup plus élevé dans les deux cas. L'avantage évident est une diminution du temps de calcul nécessaire qui est nettement inférieur aux autres méthodes simulés. Les performances sont maximales pour les paramètres $N = 512$ et $\Delta N(\%) = 50$.

Tableau 4.16

SS en fréquence, mélange instantané

	<i>N</i>						
	16	32	64	128	256	512	1024
$\Delta N(\%)$	<i>S_{IS} (dB)</i>						
12.5	2.60	2.16	4.54	5.10	5.81	4.15	3.79
25.0	2.71	2.39	4.09	5.94	5.83	5.40	4.37
37.5	2.73	2.58	4.27	4.90	5.78	5.20	2.63
50.0	2.86	2.56	4.28	4.75	5.79	6.01	4.86
62.5	2.71	2.45	4.25	5.61	4.98	5.20	3.27
75.0	2.70	2.41	4.36	4.87	5.82	5.21	3.19
	<i>D_I (dB)</i>						
12.5	-4.49	-4.21	-5.52	-5.15	-5.02	-2.32	-2.70
25.0	-4.56	-4.35	-5.23	-5.77	-5.06	-3.92	-3.23
37.5	-4.57	-4.47	-5.39	-4.91	-5.10	-3.85	-2.34
50.0	-4.65	-4.46	-5.24	-4.72	-5.11	-4.22	-2.87
62.5	-4.56	-4.39	-5.37	-5.51	-3.09	-3.86	-2.67
75.0	-4.56	-4.37	-5.43	-4.99	-5.08	-3.90	-2.68
	<i>S_{IS} (dB)</i>						
12.5	5.18	4.94	7.23	7.27	8.07	6.99	6.08
25.0	5.20	5.26	7.06	6.95	8.00	7.71	5.13
37.5	5.14	5.43	7.05	6.55	7.75	5.93	4.87
50.0	5.08	5.36	7.30	6.27	7.74	6.42	5.36
62.5	5.21	5.23	7.10	6.94	7.34	6.18	4.70
75.0	5.21	5.34	7.17	6.53	7.80	6.10	4.62
	<i>D_S (dB)</i>						
12.5	-6.22	-6.05	-7.52	-7.16	-5.50	-4.38	-3.87
25.0	-6.25	-6.29	-7.48	-7.03	-5.54	-5.06	-4.47
37.5	-6.23	-6.37	-7.47	-6.59	-5.58	-4.86	-3.70
50.0	-6.17	-6.34	-7.20	-6.34	-5.62	-5.27	-3.66
62.5	-6.25	-6.25	-7.49	-7.14	-4.32	-5.05	-4.14
75.0	-6.26	-6.33	-7.53	-6.61	-5.57	-5.07	-4.13
	<i>Temps de simulation (secondes)</i>						
12.5	7	8	6	6	7	9	10
25.0	8	8	6	7	8	9	11
37.5	10	9	10	8	9	10	13
50.0	12	12	13	10	11	12	16
62.5	15	15	17	13	14	16	19
75.0	23	23	26	20	20	23	27
<i>Diaphonie</i>				<i>Paramètres</i>			
<i>S_{II}</i>		<i>S_{II}</i>		μ		<i>n</i>	
2.54		4.85		0.10		75	

Les résultats du tableau 4.16 sont comparables à ceux obtenus avec l'utilisation de la TPO. Par contre ils sont inférieurs à ceux obtenus avec l'utilisation de la TOC. Le niveau de performance augmente avec l'augmentation du nombre de point de la TFFG. L'avantage comme dans le cas du mélange instantané est la diminution du temps de calcul nécessaire. Les performances sont maximales pour les paramètres $N = 512$ et $\Delta N(\%) = 50$.

La qualité de la séparation à l'entrée du système de séparation est présente sous deux formes soit : la qualité pleine bande S_{1i} S_{2i} et la qualité après analyse et synthèse avec la TOC S_{1ir} S_{2ir} . Ces données sont présentes pour vérifier l'impact de l'analyse et de la synthèse à lui seul sur le niveau de diaphonie. Il devient possible de vérifier l'augmentation ou la diminution engendré par la TOC et la TOCI sur la séparation.

Les tables 2 et 3 sont des compilations des données optimales obtenues avec la TOC pour chaque loi, avec la TPO et avec l'algorithme en fréquence seulement. Seul le niveau diaphonie à l'entrée après synthèse est présenté dans les tables pour la SS avec la TOC. Le niveau de diaphonie pleine bande est celui présenté pour l'algorithme de SS en fréquence et avec la TPO.

Tableau 4.17

Mesures de performances optimales, mélange instantané

Mesures de performances en dB						
Algorithme	Fréquence	TPO	TOC			
Ondelette	---	Db15	Morlet			
		Niveau 1	Uniforme	Linéaire	Log.	Mu
N	256	256	32	32	16	32
$\Delta N(\%)$	25 %	50 %	50 %	50 %	25 %	50 %
S_{ij}	2.50	2.50	3.86	3.01	2.01	2.67
S_{i2}	4.62	4.62	3.26	4.11	5.11	4.45
S_{a1}	18.00	39.39	35.46	34.41	33.37	34.54
S_{a2}	21.20	37.58	35.36	35.30	41.79	36.79
D_1	-14.13	3.82	-32.31	-31.16	-33.09	31.82
D_2	-14.89	5.16	-32.38	-32.57	-37.00	33.36

Tableau 4.18

Mesures de performances optimales, mélange convolutif

Mesures de performances en dB						
Algorithme	Fréquence	TPO	TOC			
Ondelette	---	DB15	Morlet			
		Niveau 2	Uniforme	Linéaire	Log.	Mu
N	256	64	512	512	512	512
$\Delta N(\%)$	25 %	50 %	50 %	50 %	37.5 %	37.5 %
S_{01}	2.54	2.54	4.07	3.35	3.44	3.45
S_{02}	4.85	4.85	3.57	3.97	6.40	5.08
S_{0f}	5.83	6.50	9.81	10.65	10.97	10.12
S_{02}	8.00	7.59	9.97	9.68	13.03	11.10
D_1	-5.06	-5.01	-4.60	-3.89	-5.62	-5.14
D_2	-5.54	-3.69	-5.35	-5.21	-8.98	-7.61

La figure 4.5 illustre les signaux sources, les signaux observés et les signaux sources reconstruits pour les paramètres optimaux $N = 512$ et $\Delta N(\%) = 37.5$ du tableau 4.2. Cette représentation permet de valider les résultats numériques obtenus. Les signaux sources de la figure 4.5 ne sont pas les signaux pleines bandes mais ceux obtenus après analyse et synthèse avec la TOC. La comparaison doit être faite avec ceux-ci car ce ne sont pas les signaux pleine bande que l'on retrouve dû à l'erreur de reconstruction avec la TOCI.

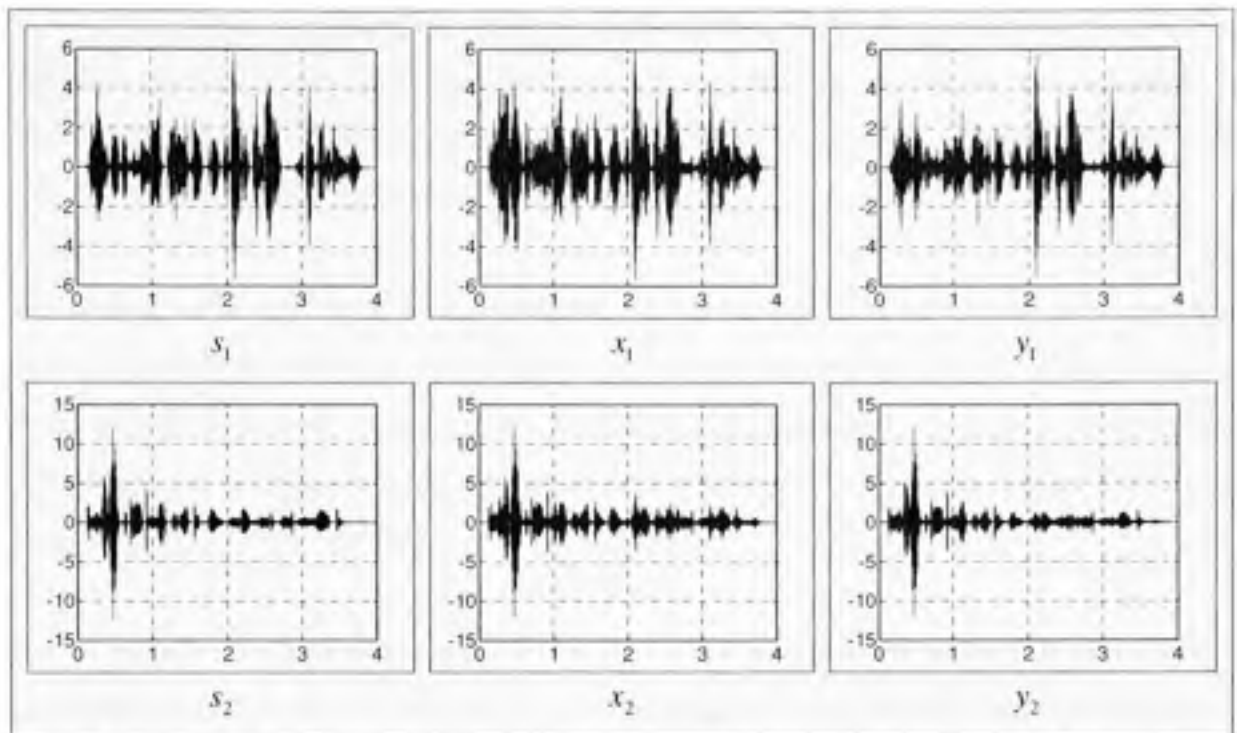


Figure 4.5 *Signaux sources, observés et reconstruits pour l'algorithme de SS avec la TOC utilisant les paramètres optimaux du tableau 4.2.*

4.7 Discussion des résultats

Les données montrent une amélioration des performances avec l'utilisation de la TPO et de la TOC comme prétraitement. Même dans le cas du mélange instantané, les performances sont améliorées.

L'amélioration peut être attribuée à plusieurs facteurs : la modification du niveau de diaphonie par la TOC, l'effet de la TOC sur la distribution statistique et le travail en sous-bandes.

L'analyse et la synthèse avec la TOC a un effet direct sur le niveau de diaphonie sans même prendre en considération le bloc de séparation de sources. Chaque échelle correspond à filtrer le signal par un filtre passe bande. Le niveau de diaphonie varie en fonction de l'échelle ou de la bande de fréquence. Le mélange des fréquences n'est pas identique sur toute la bande passante dû aux fonctions de transferts de la matrice de mélange. Le niveau

obtenu pour chaque échelle est différent de celui calculé à partir de la band complète. Cette variation du niveau de diaphonie à l'entrée du système de séparation peut causer une amélioration ou une diminution des performances. Ainsi le choix de certaines échelles a une influence sur le degré de performance de l'algorithme.

Si la distribution statistique des coefficients est altérée par la TOC les performances elles aussi s'en trouvent affectées. Si la distribution recherchée est laplacienne et que l'analyse avec la TOC pour une certaine échelle génère des coefficients ayant une distribution laplacienne les performances de séparation seront altérées. La solution risque d'être plus juste et obtenue plus rapidement.

Il a été reporté dans la littérature que le travail en sous bande améliore les performances des algorithmes de séparation. La TOC est une décomposition en sous bandes non-uniformes. Par sa propriété de travail en sous bandes il est possible d'associer l'augmentation des performances à cette affirmation.

CONCLUSION

Plusieurs approches existent pour traiter le problème que constitue la SS. Les algorithmes traitent, pour la plupart les signaux observés dans le domaine temporel ou fréquentiel. Les approches rapportées dans la littérature utilisent la TF, la décomposition en sous-bandes, la transformée de Wigner-Ville et la TO, etc.

Les algorithmes de type ondelettes utilisent la TPO et la TOD pour effectuer la SS. Les applications utilisant ces types de TO sont orientées vers le mélange instantané et convolutif.. Au meilleur de notre connaissance, aucun algorithme ne semble à ce jour, avoir été développé pour traiter un mélange convolutif de locuteurs.

Dans le cadre de ce travail une nouvelle méthode de SS a été développée pour traiter un mélange convolutif de locuteurs. L'algorithme proposé utilise la TOC comme prétraitement avant d'effectuer la séparation des coefficients en fréquence en faisant usage de l'algorithme de séparation EASI.

Les échelles choisies dans ce mémoire pour l'analyse avec la TOC sont en nombre fini et sélectionnées de façon à couvrir la bande passante utilisée en téléphonie. Les échelles sont réparties suivant quatre lois: fréquences uniformes, échelles uniformes, loi mu et loi logarithmique. Ces lois permettent entre autre de vérifier l'impact du choix des échelles sur la performance de la SS.

Les signaux observés employés pour les simulations proviennent du mélange d'un locuteur et d'une locutrice. Les dimensions du problème sont choisies pour obtenir le cas déterminé soit celui ou le nombre de sources est égal au nombre de mélanges au nombre de deux. Les fonctions de transfert utilisées pour construire les signaux observés sont synthétiques et à phases minimales. Les signaux observés obtenus sont donc synthétiques.

Les simulations sont effectuées pour trois algorithmes : SS avec la TOC, SS avec la TPO et SS en fréquence seulement. Tous ces algorithmes possèdent le même bloc de séparation en fréquence constitué par la TFFG et l'algorithme EASI. Les données obtenues permettent de comparer les performances en fonction du type de TO utilisée.

Les résultats montrent une amélioration des performances avec l'utilisation de la TOC. Cette amélioration est justifiée par la modification des distributions des signaux par la TOC et par le travail en sous-bandes. Les performances sont relativement indépendantes de la loi utilisée. Pour les quatre lois, les résultats montrent une bonne séparation des signaux observés.

L'augmentation des performances avec l'algorithme proposé peut-être attribué au travail en sous-bandes et à la modification des distributions. L'impact de la TOC est comparable à celui de la TPO lors du prétraitement. Les effets associés à l'utilisation de la TPO peuvent donc être associé à l'utilisation de la TOC.

TRAVAIL FUTUR

La première amélioration devant être apportée à l'algorithme concerne le type de mélanges pouvant être traité. L'algorithme développé est en mesure de traiter seulement les mélanges provenant d'une matrice de mélange constitué de filtres à phases minimales. Pour être complet, l'algorithme doit être en mesure de traiter les mélanges résultants d'une matrice de filtres à phases non-minimales.

L'utilisation d'un pas adaptatif, de non-linéarités avec seuil et de critères d'arrêts des itérations sont des éléments pouvant améliorer les performances de l'algorithme. Un choix automatique des échelles basé sur un critère utilisant le niveau d'information est une autre amélioration pouvant être apportée.

L'élimination des variables nécessitant un ajustement dans le but de rendre l'algorithme plus aveugle serait important. Les paramètres les plus dérangeants sont actuellement la longueur et le déplacement avec la TFFG ainsi que le nombre et la répartition des échelles avec la TOC. L'élimination de tous ces paramètres rendrait les performances de l'algorithme dépendant seulement de la non-linéarité et de l'ondelette choisie.

Des tests de robustesses de l'algorithme par rapport aux mélanges devraient être développés avec le développement d'un système en temps réel. La transformation de cet algorithme en un algorithme en temps réel est une avenue intéressante. Cette implantation permettrait vraiment d'évaluer les capacités de cette approche à effectuer la SS.

BIBLIOGRAPHIE

1. Haykin, S., *Unsupervised adaptive filtering*. 2000, New York: J. Wiley and Sons. 2 v.
2. Lacoume, J.L., P.O. Amblard, and P. Comon, *Statistiques d'ordre supérieur pour le traitement du signal*. Traitement du signal. 1997, Paris: Masson. xiv, 290.
3. Schobben, D.W.E., *Real-time adaptive concepts in acoustics : blind signal separation and multichannel echo cancellation*. 2001, Dordrecht ; Boston: Kluwer Academic Publishers. xii, 160.
4. Molgedey, L. and G. Schuster, *Separation of a mixture of independent signals using time delayed correlations*. Physical Review Letters, 1994. **vol. 72**(no. 23): p. 3634--3637.
5. Ananthanag, K.V.K. and J.S. Sahambi. *Investigation of blind source separation methods for extraction of fetal ECG*. in *IEEE Canadian Conference on Electrical and Computer Engineering CCECE*. 2003.
6. Diamantaras, K.I. and T. Papadimitriou. *Blind separation of reflections using the image mixtures ratio*. in *IEEE International Conference on Image Processing, 2005. ICIP 2005*. 2005.
7. Ehlers, F. and H.G. Schuster, *Blind separation of convolutive mixtures and an application in automatic speech recognition in a noisy environment*. IEEE Transactions on Signal Processing, 1997. **45**(10): p. 2608-2612.
8. Escudero, J., et al., *Artifact Removal in Magnetoencephalogram Background Activity With Independent Component Analysis*. IEEE Transactions on Biomedical Engineering. **54**(11): p. 1965-1973.
9. Hyvärinen, A. and E. Oja, *Independent Component Analysis: Algorithms and Applications*. Helsinki University of Technology Neural Networks Research Centre, 2000.
10. Jafari, M.G. and J.A. Chambers, *Fetal electrocardiogram extraction by sequential source separation in the wavelet domain*. IEEE Transactions on Biomedical Engineering, 2005. **52**(3): p. 390-400.
11. Kam, A. and A. Cohen. *Separation of twins fetal ECG by means of blind source separation (BSS)*. in *The 21st IEEE Convention of the Electrical and Electronic Engineers in Israel*. 2000.
12. Lambert, R.H., *Multichannel blind deconvolution: FIR matrix algebra and separation of multipath mixture*. 1996, University of Southern California.

13. Mitianoudis, N. and M.E. Davies, *Audio source separation of convolutive mixtures*. IEEE Transactions on Speech and Audio Processing, 2003. **11**(5): p. 489-497.
14. Oja, E., K. Kiviluoto, and S. Malaroiu, *Independent component analysis for financial time series*. in *Symposium IEEE Adaptive Systems for Signal Processing, Communications, and Control AS-SPCC*. 2000.
15. Tonazzini, A., I. Gerace, and F. Cricco, *Joint blind separation and restoration of mixed degraded images for document analysis*. in *International Conference on Image Processing ICIP*, 2004.
16. Vigario, R., et al., *Independent component approach to the analysis of EEG and MEG recordings*. IEEE Transactions on Biomedical Engineering, 2000. **47**(5): p. 589-593.
17. Xiaopei, W. and G. Xiaojing, *Mental EEG analysis based on independent component analysis*. in *Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis ISPA*. 2003.
18. Zhou, J., et al. *Subtraction of Ocular and Line Artifacts from Mental EEG based on ICA*. in *Sixth International Conference on Intelligent Systems Design and Applications ISDA*. 2006.
19. Arons, B., *A Review of the Cocktail Party Effect*. MIT Media Lab, 1992.
20. Cherryl, E.C., *Some experiments on the recognition of speech, with one and two ears*. Journal of Acoustic Society of America, 1953. **25**: p. 975-979.
21. Jutten, C., J. Herault, and A. Guerin, *IIN.C.A: An independent components analyzer based on an adaptive neuromimetic network*. in: J.Degmongeot, T.Herve, V.Raille and C.Roche, eds., *Artificial Intelligence and Cognitive Science*, Manhester Univ. Press, Manchester, 1988.
22. Cardoso, J.F. and A. Souloumiac, *Blind beamforming for non gaussian signals*. IEE-Proceedings-F, 1993. **vol. 140, no. 6**: p. 362-370.
23. Cardoso, J.F. and B.H. Laheld, *Equivariant adaptive source separation*. IEEE Transactions on Signal Processing, 1996. **44**(12): p. 3017-3030.
24. Hyvarinen, A., *Fast and robust fixed-point algorithms for independent component analysis*, Neural Networks, IEEE Transactions on, 1999. **10**(3): p. 626-634.
25. Bell, A.J. and T.J. Sejnowski, *An information-maximisation approach to blind separation and blind deconvolution*. Neural Computation, 1995. **vol. 7**: p. 1129-1159.
26. Murata, N. and S. Ikeda, *An approach to blind source separation based on temporal structure of speech signals*. RIKEN Brain Science Institute, 1998.

27. Bose, T., *Digital signal and image processing*, 2004, Hoboken, N.J.: J. Wiley. xiv, 706.
28. Févrotte, C. and C. Doncarli, *A unified presentation of blind signal separation method for convolutive mixtures using block-diagonalisation*. 4th International Symposium on Independent Component Analysis and Blind Signal Separation, Nara, Japan, April 2003.
29. Cardoso, J.F., *Separation of non stationary sources: Achievable performance*. In Proc. SSAP, 2000: p. 359-363.
30. Cardoso, J.-F., *Blind Signal Separation: Statistical Principles*. Proceedings of the IEEE, 1998. **86**(10): p. 2009-2025.
31. Leon, P.D. *Impulse Responses for Various Acoustic Environments*. 2007 [cited 2007 4 mars]; Available from: <http://www.ece.nmsu.edu/~pdeleon/ASAP/BSS/>.
32. Lee, T.W., A.J. Bell, and R. Orglmeister, *Blind source separation of real world signals*. in *International Conference on Neural Networks*. 1997.
33. Zhang, L.Q., A. Cichocki, and S. Amari, *Multichannel Blind Deconvolution of Non-minimum Phase Systems Using Information Backpropagation*. In Proceedings of 6th International Conference on Neural Information Processing ICONIP, 1999: p. 210-216.
34. Pham, D.T. and J.F. Cardoso, *Blind separation of instantaneous mixtures of nonstationary sources*. IEEE Transactions on Signal Processing, 2001. **49**(9): p. 1837-1848.
35. Gazor, S. and W. Zhang, *Speech Probability Distribution*. Signal Processing Letters, IEEE, 2003. **10**(7): p. 204-207.
36. Suetake, N., Y. Nakamura, and T. Yamakawa. *Maximum Entropy ICA Constrained by Individual Entropy Maximization Employing Self-organizing Maps*. in *International Joint Conference on Neural Networks IJCNN*. 1999.
37. Hyvarinen, A., *Fast and robust fixed-point algorithms for independent component analysis*. Neural Networks, IEEE Transactions on, 1999 **10**(3): p. 626-634.
38. Cardoso, J.-F. and T. Adali. *The Maximum Likelihood Approach to Complex ICA*. in *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP*, 2006.
39. Aichnery, R., et al. *Time domain blind source separation of non-stationary convolved signals by utilizing geometric beamforming*. in *Proceedings of the 2002 12th IEEE Workshop on Neural Networks for Signal Processing*. 2002.

40. Theis, F.J., A. Mayer-Baese, and W.E. Lang, *Second-order blind source separation based on multi-dimensional autocovariances*. University of Regensburg, Florida State University, 2004.
41. Winter, S., et al., *Underdetermined source separation by ICA and homomorphic signal processing*. IWAENC – PARIS, 2006: p. 12-14.
42. Takada, K., et al., *On Wavelet-based Convolutional Blind Separation of Non-Stationary Sound Sources*. SICE Annual Conference in Sapporo, 2004: p. 1362-1367.
43. Jafari, M.G. and J.A. Chambers, *Wavelet domain natural gradient algorithm for blind source separation of non-stationary sources*. Electronics Letters, 2002. **38**(14): p. 759-761.
44. Fahmy, M.F., et al., *A Fast Wavelet Packet Based Blind Signal Separation*. 21st National Radio Science Conference (NRSC), 2004.
45. Hirai, N., et al. *A consideration of blind source separation using wavelet transform*. in *Proceedings of International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*. 2004.
46. Moussaoui, R., J. Rouat, and R. Lefebvre. *Wavelet Based Independent Component Analysis for Multi-Channel Source Separation*. in *ICASSP Proceedings. IEEE International Conference on Acoustics, Speech and Signal Processing* 2006.
47. Fasong, W., L. Hongwei, and L. Rui. *Applications of Wavelet to Independent Component Analysis*. in *The Sixth World Congress on Intelligent Control and Automation WCICA*. 2006.
48. Russel, I., et al., *Integration of DFT and cosine-modulated filter banks with blind separation of convolutively mixed non-stationary sources*. University of Wollongon, Faculty of Informatics, 2004.
49. Duplessis-Beaulieu, F., *Fast Convolutional Blind Speech Separation via Subband Adaptation*, in *Department of Electrical and Computer Engineering*. 2002, McGill University: Montreal.
50. Weinstein, E., M. Feder, and A.V. Oppenheim, *Multi-channel signal separation by decorrelation*. IEEE Transactions on Speech and Audio Processing, 1993. **1**(4): p. 405-413.
51. Polikar, R. *The Wavelet Tutorial*. 1999 [cited 2007 2007]; Available from: <http://users.rowan.edu/~polikar/WAVELETS/WTtutorial.html>.
52. Mitra, S.K., *Digital signal processing : a computer based approach*. 3rd ed. 2006, New York, NY: McGraw-Hill Higher Education. xx, 972.

53. Vaidyanathan, P.P., *Multirate systems and filter banks*. Prentice-Hall signal processing series. 1993, Englewood Cliffs, N.J.: PTR Prentice Hall. xvi, 911.
54. Strang, G. and T. Nguyen, *Wavelets and filter banks*. 1996, Wellesley, Mass.: Wellesley-Cambridge Press. xxi, 490.
55. Kronland-Martinet, R., *The Wavelet Transform for Analysis, Synthesis and Processing of Speech and Music Sounds*. Computer Music Journal, 1988. Vol. 12(no. 4): p. 11-20.
56. Sarkar, T., M. Salazar-Palma, and M.C. Wicks, *Wavelet applications in engineering electromagnetics*. Artech House electromagnetic analysis series. 2002, Boston: Artech House. xxi, 347.
57. De Gersem, P.D., B.D. De Moor, and M. Moonen. *Applications of the continuous wavelet transform in the processing of musical signals*. 1997.
58. Goswami, J.C. and A.K. Chan, *Fundamentals of wavelets : theory, algorithms, and applications*. Wiley series in microwave and optical engineering. 1999, New York, N.Y.: J. Wiley and Sons. xvi, 306.
59. Mallat, S.G., *A wavelet tour of signal processing*. 1998, New York: Academic Press. xxii, 577.
60. Akansu, A.N. and R.A. Haddad, *Multiresolution signal decomposition : transforms, subbands, and wavelets*. Telecommunications. 1992, San Diego, Calif.: Academic Press. xii, 376.
61. MATLAB, *Wavelet Packet Analysis, Wavelets: A new tool for signal analysis (Wavelet Toolbox)*. 2004.
62. Akansu, A.N., Smith, and J.T. Mark, *Subband and wavelet transforms design and applications*. Kluwer international series in engineering and computer science. 1996, Boston, Mass.: Kluwer Academic. xvi, 451.
63. Hubbard, B.B., *Ondes et ondelettes : la saga d'un outil mathématique*. 1995, Paris: Pour la Science diffusion Belin. 236.
64. Teolis, A., *Computational signal processing with wavelets*. 1998, Boston, Mass.: Birkhäuser. xxiv, 324.
65. Saha, S. *Image Compression - from DCT to Wavelets : A Review*. 2000 [cited; Available from: <http://www.acm.org/crossroads/xrds6-3/sahaimgcoding.html>].
66. Lazar, D. and A. Averbuch, *Wavelet-based video coder via bit allocation*. IEEE Transactions on Circuits and Systems for Video Technology, 2001. 11(7): p. 815-832.

67. Chen, G.Y. and B. Kegl, *Invariant Pattern Recognition using Dual-tree Complex Wavelets and Fourier Features*. Department of Computer Science and Operations Research, University of Montreal,.
68. Belley, L.P., M. Gabrea, and C. Gargour, *Continuous wavelet transform based source separation*, in *IEEE International Midwest Symposium on Circuits and Systems MWSCAS. Digital and adaptive signal processing*. 2007: Montreal.
69. Mathis, H., T.P. Hoff, and M. Joho, *Blind Separation of Signals with Mixed Kurtosis Signs Using Threshold Activation Functions*. *IEEE Transactions on Neural Networks*,, 2001. **12**(3): p. 618-624.
70. TIMIT, *Training and Test Data and Speech Header Software*. 1990, The DARPA TIMIT Acoustic-Phonetic Continuous Speech Corpus.
71. McGovern, S.G. *A Model for Room Acoustics*. 2003-2004 [cited; Available from: <http://www.2pi.us/rir.html>].
72. Schobben, D., K. Torkkola, and P. Smaragdis, *Evaluation of Blind Signal Separation Methods*. 1999.
73. Cheikhrouhou, I., et al. *Speech Analysis Using Wavelet Transforms Dedicated To Cochlear Prosthesis Stimulation Strategy*. in *Control, Communications and Signal Processing. First International Symposium on*. 2004.
74. T. Tan, B., et al. *The Use Of Wavelet Transforms In Phoneme Recognition*. in *Fourth International Conference on Spoken Language Proceedings ICSLP*. 1996.
75. Tian, G., Y. Shuli, and Y. Datian. *Application of Wavelet in Speech Processing of Cochlear Implant*. in *IEEE-EMBS 27th Annual International Conference of the Engineering in Medicine and Biology Society*. 2005.
76. Agbinya, J.I. *Discrete wavelet transform techniques in speech processing*. in *TENCON '96. Proceedings. IEEE TENCON. Digital Signal Processing Applications*. 1996.
77. Gamulkiewicz, B. and M. Weeks. *Wavelet based speech recognition*. in *Proceedings of the 46th IEEE International Midwest Symposium on Circuits and Systems MWSCAS*. 2003.
78. Xueying, Z. and J. Zhiping. *Speech recognition based on auditory wavelet packet filter*. in *Proceedings of the 7th International Conference on Signal Processing ICSP* 2004.